

Study on the azimuthal angle correlation between two jets in the top quark pair production

Kaoru Hagiwara¹ and Junya Nakamura²

^{1,2}*KEK Theory Center and Sokendai,
Tsukuba, Ibaraki 305-0801, Japan*

²*Institute for Theoretical Physics, University of Tübingen,
Auf der Morgenstelle 14, 72076 Tübingen, Germany.*

Abstract

An azimuthal angle correlation between the two hardest jets is studied in the $t\bar{t}$ production process at the 14 TeV LHC. The event samples are generated by merging the tree level matrix elements for the $t\bar{t}$ plus up to 2 or 3 partons with parton showers. The generated event samples show a strong correlation in the azimuthal angle difference between the two hardest jets, as predicted in the analysis based on the tree level matrix elements for the $t\bar{t} + 2$ partons. The effects of merging the matrix elements for the $t\bar{t} + 3$ partons on the correlation are studied in detail. It is found that they play important roles in improving the prediction of the correlation.

¹E-mail: kaoru.hagiwara@kek.jp

²E-mail: junya.nakamura@uni-tuebingen.de

Contents

1	Introduction	3
2	Implementation of merging algorithms	4
2.1	Tree level merging algorithms	5
2.2	The CKKW-L algorithm and its extension	8
2.3	Construction of the PYTHIA8 parton shower history	12
2.3.1	Construction of the ISR history	12
2.3.2	Construction of the FSR history	13
2.3.3	Some technical details	14
2.4	Weight functions	15
2.5	The merging algorithm with top decays	17
2.6	Event generation	18
2.7	Tests of our implementation	20
3	Azimuthal angle correlation	22
3.1	Definition	23
3.2	Contamination	23
3.3	The effects of the $t\bar{t}+3$ partons matrix elements	25
3.4	The effect of a $t\bar{t}$ decay	28
4	Conclusion	29

1 Introduction

Since the discovery of the Higgs boson was announced in the summer of 2012, the LHC measurements of its properties have so far been supporting the standard model (SM) predictions [1, 2, 3, 4]. The Higgs sector of the SM respects the charge-conjugation and parity (CP) symmetry and the Higgs boson should be CP even. Therefore if an admixture of the CP odd component is observed, it will be a direct evidence of CP violation in the Higgs sector and thus physics beyond the SM.

From the analyses on the tree level matrix elements, it has been shown that the azimuthal angle difference between the two partons (gluon, quarks or antiquarks) produced in association with the Higgs boson produced by gluon fusion is very sensitive to the CP property of the Higgs boson [5, 6, 7, 8]. Several analyses including effects of higher order corrections show that the correlation between the two partons found at the tree level matrix elements can be observed as a correlation between the two hardest *jets* despite smearing, see e.g. refs. [9, 10, 11, 12]. However the attempts to observe the CP odd admixture precisely in this approach are expected to be difficult due to large theoretical uncertainties in Monte Carlo event simulation, particularly in our use of a parton shower generator which can simulate the QCD radiation only in the soft and/or collinear limit.

It has been pointed out in ref. [13] that the two partons produced in association with a top quark pair has a large azimuthal angle correlation near the threshold $m_{t\bar{t}} \sim 2m_t$ and the correlation is similar to that of the two partons produced together with the CP odd Higgs boson via gluon fusion. The claim of ref. [13] is that experimental techniques to measure such an angular correlation between jets can be established first by using these SM processes which have large cross sections. More precisely, we measure the azimuthal angle difference between two jets produced in association with a top quark pair and tune a Monte Carlo event generator to reproduce the data quantitatively. If an event generator tuned in this way is used, the theoretical uncertainty on the prediction of the azimuthal angle correlation between two jets produced in association with the Higgs boson can be reduced significantly. This will help achieve accurate measurements of the CP property of the Higgs boson.

In the present paper we attempt to create a bridge between the proposal of ref. [13] and actual experimental measurements, by studying our present capability and limitation of simulating the top quark pair plus multi-jet production process, so that experimentalists can use the real data to improve our simulation tools to be used to probe more fundamental physics such as the CP property of the Higgs boson. The simplest method to include leading higher order corrections to the top quark pair production is to apply a parton shower generator to the exclusive top quark pair events, where the parton shower scale evolution of the top quark pair events produces the top quark pair plus multi-jet events at a hadronization scale. The event samples generated in this way are expected to reproduce qualitatively the multi-jet event rates and the jet p_T and rapidity distribution, since the successive emission of parton showers follows the QCD prediction in the soft and/or collinear region and the overall jet rates have been fitted to the data in e^+e^- and hadronic collisions. Those events, however, do not have correct correlations among jets since a parton shower generator emits azimuthally symmetric radiation about a parent momentum direction. To reproduce azimuthal angle correlations between two jets, at least the $t\bar{t} + 2$ partons matrix elements have to be embedded. In order to consistently combine the event samples for different parton multiplicity generated by tree level matrix elements with parton showers, a tree level merging algorithm is required. There exists several tree level merging algorithms proposed in literatures, including the CKKW [14, 15, 16, 17], the CKKW-L [18, 19, 20], the GKS [21], the MLM [22, 23], the pseudo shower [24] and the shower k_\perp [25] algorithms. Comparisons between these different algorithms have also been studied [24, 25, 26, 27]. Our objective in the present paper is to implement tree level merging algorithms and study theoretical issues on predicting the azimuthal angle correlation between the two hardest jets.

We generate the event samples for the top quark pair production at the 14 TeV LHC, by merging the tree level matrix elements for the $t\bar{t}$ plus up to 2 or 3 partons with the parton shower model in PYTHIA8. By using the generated event samples, the azimuthal angle difference between the two hardest jets (i.e. the two highest transverse momentum jets), $\Delta\phi = \phi_1 - \phi_2$, is studied.

As tree level merging algorithms, the CKKW-L merging algorithm [18, 19, 20] and a new tree level merging algorithm are implemented. Our new algorithm differs from the CKKW-L algorithm in the strategy for phase space separation. It is designed so that the contribution from the $t\bar{t} + 0, 1$ parton matrix elements to the event samples with two or more jets, which we call the contamination, can be more suppressed above the merging scale. Therefore, more accurate predictions on correlations between two jets are expected. We confirm this by numerically comparing the two algorithms. The contamination is studied by varying a relation between the merging scale and the scale of a jet definition. We find that the contamination is not negligible when the merging scale is set equal to or slightly smaller than the scale of the anti- k_T jet definition.

We produce the $\Delta\phi$ distribution by using the generated event samples. The distribution shows a strong correlation in $\Delta\phi$, as predicted in the previous analysis [13] based on the $t\bar{t} + 2$ partons tree level matrix elements. This observation confirms that the correlation found in the $t\bar{t} + 2$ partons tree level matrix elements [13] is still visible after including the dominant QCD higher order corrections and thus can be observed in the experiments.

We observe a clear difference in the $\Delta\phi$ distribution between the event samples generated by merging the matrix elements for the $t\bar{t}$ plus up to 2 partons and those generated by merging the matrix elements for the $t\bar{t}$ plus up to 3 partons. Furthermore, the difference is found slightly larger, when the rapidity range for jets is more restricted. We study the effects of the $t\bar{t} + 3$ partons matrix elements on the correlation in detail and find out the origins of the difference. We show that the $t\bar{t} + 3$ partons matrix elements play important roles in predicting $\Delta\phi$ accurately.

We present a method for merging the matrix element event samples which include a $t\bar{t}$ decay as a part of the hard process with the parton shower. In this method, correlations between the decay products of the $t\bar{t}$ are predicted correctly, while a merging algorithm is performed as if the $t\bar{t}$ was not decayed¹. The effect of the $t\bar{t}$ dilepton decay on the azimuthal angle correlation is studied. We find that the effect is small, when the two hardest jets are picked up from all jets not including the two hardest b jets.

We note in passing that this is not the first attempt to estimate higher order corrections to the azimuthal angle difference between the two partons produced in association with a top quark pair. The azimuthal angle difference between two jets in the $t\bar{t}$ production process has been studied with the aim of a scalar top quark search in ref. [28], of a gluino search in ref. [29] and of investigating top quark mass effects in the effective Higgs-gluon coupling in refs. [30] and [31].

In Section 2, our implementation of the tree level merging algorithms is described in detail. In Section 3, the azimuthal angle correlation is studied. In Section 4, we summarize our findings.

2 Implementation of merging algorithms

Our implementation of tree level merging algorithms is described in this section. In Section 2.1, the basic idea of tree level merging algorithms [14] is reviewed. Our notations used throughout the paper are also introduced. In Section 2.2, the CKKW-L merging algorithm [18, 19, 20] is reviewed at first. After that, we introduce a new tree level merging algorithm, which differs from the CKKW-L merging algorithm in the strategy for phase space separation. The procedure of constructing the PYTHIA8 parton shower history is described in Section 2.3 and that of calculating the weight function is explained in Section 2.4. These are ingredients of the merging algorithms. A method for consistently merging the matrix elements event samples which include decays of the top and antitop quarks as a part of the hard process is presented in Section 2.5. In Section 2.6, a procedure for the event generation of the top quark pair production and setups for it such as scale choices are explained. Our implementation is carefully tested in Section 2.7.

¹This is necessary, since QCD radiation off a top quark takes place faster than its decay.

2.1 Tree level merging algorithms

In this section, the basic idea of tree level merging algorithms [14] is reviewed. Our notations used throughout the paper are also presented.

Let us start with the DGLAP evolution equation [32, 33, 34] with the Sudakov form factor [35]

$$t \frac{d}{dt} \frac{q(x, t)}{\Delta(t)} = \int_0^{\epsilon(t)} \frac{dz}{z} \frac{\alpha_s}{2\pi} \hat{P}_{qq}(z) \frac{q(x/z, t)}{\Delta(t)}, \quad (2.1)$$

where the Sudakov form factor is given by

$$\Delta(t) = \exp \left(- \int_{\mu^2}^t \frac{dt'}{t'} \int_0^{\epsilon(t')} dz \frac{\alpha_s}{2\pi} \hat{P}_{qq}(z) \right). \quad (2.2)$$

Here only the quark parton distribution function $q(x, t)$ (PDF) and the splitting function without the virtual correction $\hat{P}_{qq}(z)$ for $q \rightarrow qq$ are introduced in order to simplify our writing. The generalization is, however, simple. Integrating eq. (2.1) over $t_\Lambda < t < t_X$ gives

$$\frac{q(x, t_X)}{\Delta(t_X)} = \frac{q(x, t_\Lambda)}{\Delta(t_\Lambda)} + \int_{t_\Lambda}^{t_X} \frac{dt}{t} \int_0^{\epsilon(t)} d\hat{p}_{qq}(z) \frac{q(x/z, t)}{\Delta(t)}, \quad (2.3)$$

where a short hand notation is used

$$d\hat{p}_{qq}(z) = \frac{\alpha_s}{2\pi} \frac{dz}{z} \hat{P}_{qq}(z). \quad (2.4)$$

After the infinite iterations of eq. (2.3), it is found that

$$\begin{aligned} \frac{q(x, t_X)}{\Delta(t_X)} &= \frac{q(x, t_\Lambda)}{\Delta(t_\Lambda)} + \int_{t_\Lambda}^{t_X} \frac{dt_1}{t_1} \int_0^{\epsilon(t_1)} d\hat{p}_{qq}(z_1) \frac{q(x/z_1, t_\Lambda)}{\Delta(t_\Lambda)} \\ &+ \int_{t_\Lambda}^{t_X} \frac{dt_1}{t_1} \int_0^{\epsilon(t_1)} d\hat{p}_{qq}(z_1) \int_{t_\Lambda}^{t_1} \frac{dt_2}{t_2} \int_0^{\epsilon(t_2)} d\hat{p}_{qq}(z_2) \frac{q(x/(z_1 z_2), t_\Lambda)}{\Delta(t_\Lambda)} + \dots \end{aligned} \quad (2.5)$$

By dividing this equation by $q(x, t_X)/\Delta(t_X)$, we find [35]

$$\begin{aligned} 1 &= \frac{q(x, t_\Lambda)}{q(x, t_X)} \frac{\Delta(t_X)}{\Delta(t_\Lambda)} + \int_{t_\Lambda}^{t_X} \frac{dt_1}{t_1} \int_0^{\epsilon(t_1)} d\hat{p}_{qq}(z_1) \frac{q(x/z_1, t_\Lambda)}{q(x, t_X)} \frac{\Delta(t_X)}{\Delta(t_\Lambda)} \\ &+ \int_{t_\Lambda}^{t_X} \frac{dt_1}{t_1} \int_0^{\epsilon(t_1)} d\hat{p}_{qq}(z_1) \int_{t_\Lambda}^{t_1} \frac{dt_2}{t_2} \int_0^{\epsilon(t_2)} d\hat{p}_{qq}(z_2) \frac{q(x/(z_1 z_2), t_\Lambda)}{q(x, t_X)} \frac{\Delta(t_X)}{\Delta(t_\Lambda)} + \dots \end{aligned} \quad (2.6)$$

The first term of the right hand side (RHS) in this equation can be regarded as the probability of generating no radiation from a quark $q(x)$ in a proton during the scale evolution of the proton between t_X and t_Λ ($t_X > t_\Lambda$). The second term represents the integrated probability of generating exclusively one radiation from the quark during the scale evolution, and so on. The left hand side (LHS) ensures the probability conservation.

We write the no radiation probability from a quark $q(x)$ in a proton during the scale evolution of the proton between t_1 and t_2 ($t_1 > t_2$) in the following form [35]

$$\Pi_q(t_1, t_2; x) = \frac{q(x, t_2)}{q(x, t_1)} \frac{\Delta(t_1)}{\Delta(t_2)}. \quad (2.7)$$

Then, we can express eq. (2.6) as

$$\begin{aligned}
1 = & \Pi_q(t_X, t_\Lambda; x) + \int_{t_\Lambda}^{t_X} \frac{dt_1}{t_1} \int_0^{\epsilon(t_1)} d\hat{p}_{qq}(z_1) \Pi_q(t_X, t_1; x) \frac{q(x/z_1, t_1)}{q(x, t_1)} \Pi_q(t_1, t_\Lambda; x/z_1) \\
& + \int_{t_\Lambda}^{t_X} \frac{dt_1}{t_1} \int_0^{\epsilon(t_1)} d\hat{p}_{qq}(z_1) \int_{t_\Lambda}^{t_1} \frac{dt_2}{t_2} \int_0^{\epsilon(t_2)} d\hat{p}_{qq}(z_2) \Pi_q(t_X, t_1; x) \frac{q(x/z_1, t_1)}{q(x, t_1)} \Pi_q(t_1, t_2; x/z_1) \\
& \times \frac{q(x/(z_1 z_2), t_2)}{q(x/z_1, t_2)} \Pi_q(t_2, t_\Lambda; x/(z_1 z_2)) \\
& + \dots
\end{aligned} \tag{2.8}$$

It is an easy task to guess the explicit form of $\Pi_q(t_1, t_2; x)$ from the above equation [36],

$$\Pi_q(t_1, t_2; x) = \exp \left(- \int_{t_2}^{t_1} \frac{dt}{t} \int_0^{\epsilon(t)} d\hat{p}_{qq}(z) \frac{q(x/z, t)}{q(x, t)} \right). \tag{2.9}$$

Given the scale t_X and the energy fraction x of a quark in a proton, eq. (2.8) allows us to generate radiations from the incoming quark by evolving the proton from the scale t_X . This is known as backward evolution [36, 37, 35].

The above equation in eq. (2.8) derived from the DGLAP equation concerns only radiation from an incoming quark in a proton i.e. initial state radiation. Here we generalize the equation to the one which can predict radiation from outgoing partons i.e. final state radiation as well as initial state radiation. What we should notice for this purpose is that the PDFs play a role in constraining scale evolution of the proton, or in other words constraining radiation from the quark in the proton during scale evolution of the proton. Hence, when radiation from outgoing partons is considered, we replace the PDFs with a function which is obtained from the kinematic information of the outgoing partons. The function constrains radiation from the outgoing partons, through the energy and momentum conservation for instance.

We let $\{p\}_{X+n}$ denotes a complete specification of an event sample consisting of $X + n$ partons². The information of the two incoming partons is also included. Then, we introduce a function for the evolution of a $\{p\}_{X+n}$ as

$$f(z, t; \{p\}_{X+n}), \tag{2.10}$$

which constrains the evolution of the $\{p\}_{X+n}$ at the evolution scale t and the energy fraction z . With the constraint function, eq. (2.8) can be generalized to

$$\begin{aligned}
1 = & \Pi(t_X, t_\Lambda; \{p\}_X) + \int_{t_\Lambda}^{t_X} \frac{dt_1}{t_1} \int_0^1 d\hat{p}(z_1) \Pi(t_X, t_1; \{p\}_X) f(z_1, t_1; \{p\}_X) \Pi(t_1, t_\Lambda; \{p\}_{X+1}) \\
& + \int_{t_\Lambda}^{t_X} \frac{dt_1}{t_1} \int_0^1 d\hat{p}(z_1) \Pi(t_X, t_1; \{p\}_X) f(z_1, t_1; \{p\}_X) \\
& \times \int_{t_\Lambda}^{t_1} \frac{dt_2}{t_2} \int_0^1 d\hat{p}(z_2) \Pi(t_1, t_2; \{p\}_{X+1}) f(z_2, t_2; \{p\}_{X+1}) \Pi(t_2, t_\Lambda; \{p\}_{X+2}) \\
& + \dots,
\end{aligned} \tag{2.11}$$

and accordingly

$$\Pi(t_1, t_2; \{p\}_{X+n}) = \exp \left(- \int_{t_2}^{t_1} \frac{dt}{t} \int_0^1 d\hat{p}(z) f(z, t; \{p\}_{X+n}) \right), \tag{2.12}$$

²This expression is inspired by refs. [21, 38].

which is defined as the no radiation probability for a $\{p\}_{X+n}$ as a whole, during the scale evolution of it between t_1 and t_2 ($t_1 > t_2$). In other words, this is the probability that a $\{p\}_{X+n}$ remains the same during the scale evolution. Note that

$$\begin{aligned} d\hat{p}(z) &= \frac{dz}{z} \frac{\alpha_s}{2\pi} \hat{P}(z) \quad \text{for initial state radiation,} \\ d\hat{p}(z) &= dz \frac{\alpha_s}{2\pi} \hat{P}(z) \quad \text{for final state radiation,} \end{aligned} \quad (2.13)$$

where the appropriate splitting function(s) should be used for $\hat{P}(z)$ according to a branching process $\{p\}_{X+n} \rightarrow \{p\}_{X+n+1}$. For initial state radiation, the constraint function $f(z, t; \{p\}_{X+n})$ always includes the PDFs. The soft gluon singularity at $z = 1$ in the splitting functions will be avoided by introducing θ functions in the constraint function.

Now that we have the integrated form of the DGLAP equation with the Sudakov form factor in our notations as eqs. (2.11) (2.12), we discuss the basic idea of tree level merging algorithms [14]. Let us first define the cross section of a hard process producing X by $\sigma(X)$. For example, when $X = q\bar{q}$ in e^+e^- annihilation is considered, $\sigma(X)$ is given by

$$\sigma(e^+e^- \rightarrow q\bar{q}) = \frac{1}{2s} \int d\Phi_{q\bar{q}} \sum_{\lambda} |\overline{\mathcal{M}}_{e^+e^- \rightarrow q\bar{q}}|^2, \quad (2.14)$$

or, when $X = Z$ in proton proton collisions is considered, $\sigma(X)$ is given by

$$\sigma(pp \rightarrow Z) = \int_0^1 dx_1 \int_0^1 dx_2 q(x_1, \mu_F^2) \bar{q}(x_2, \mu_F^2) \frac{1}{2sx_1x_2} \int d\Phi_Z \sum_{\lambda} |\overline{\mathcal{M}}_{q\bar{q} \rightarrow Z}|^2. \quad (2.15)$$

The DGLAP evolution of the hard process can be expressed by multiplying $\sigma(X)$ by eq. (2.11), namely

$$\begin{aligned} \sigma(X) &= \sigma(X) \Pi(t_X, t_\Lambda; \{p\}_X) \\ &+ \sigma(X) \int_{t_\Lambda}^{t_X} \frac{dt_1}{t_1} \int_0^1 d\hat{p}(z_1) \Pi(t_X, t_1; \{p\}_X) f(z_1, t_1; \{p\}_X) \Pi(t_1, t_\Lambda; \{p\}_{X+1}) \\ &+ \sigma(X) \int_{t_\Lambda}^{t_X} \frac{dt_1}{t_1} \int_0^1 d\hat{p}(z_1) \Pi(t_X, t_1; \{p\}_X) f(z_1, t_1; \{p\}_X) \\ &\quad \times \int_{t_\Lambda}^{t_1} \frac{dt_2}{t_2} \int_0^1 d\hat{p}(z_2) \Pi(t_1, t_2; \{p\}_{X+1}) f(z_2, t_2; \{p\}_{X+1}) \Pi(t_2, t_\Lambda; \{p\}_{X+2}) \\ &+ \dots \end{aligned} \quad (2.16)$$

The core idea of merging algorithms is to replace the terms constructed by the leading order cross section times the universal radiation probability with the exact tree level cross sections. For the second term of the RHS in the above equation, for instance, it proceeds as

$$\sigma(X) \int_{t_\Lambda}^{t_X} \frac{dt_1}{t_1} \int_0^1 d\hat{p}(z_1) \rightarrow \sigma(X+1). \quad (2.17)$$

The equation in eq. (2.16) has the following expression

$$\begin{aligned} \sigma(X) &= \sigma(X) \Pi(t_X, t_{\text{cut}}; \{p\}_X) \\ &+ \sigma(X+1; \{p\}_{X+1} > t_{\text{cut}}) \Pi(t_X, t_1; \{p\}_X) f'(z_1, t_1; \{p\}_X) \Pi(t_1, t_{\text{cut}}; \{p\}_{X+1}) \\ &+ \sigma(X+2; \{p\}_{X+2} > t_{\text{cut}}) \Pi(t_X, t_1; \{p\}_X) f'(z_1, t_1; \{p\}_X) \Pi(t_1, t_2; \{p\}_{X+1}) \\ &\quad \times f'(z_2, t_2; \{p\}_{X+1}) \Pi(t_2, t_{\text{cut}}; \{p\}_{X+2}) \\ &+ \dots, \end{aligned} \quad (2.18)$$

where t_Λ is replaced with t_{cut} . The soft and collinear divergences in tree level matrix elements are regularized in the following way. By using the definition of the evolution variable t , calculate the minimum value t_{min} from $\{p\}_{X+n}$ and require $t_{\text{min}} > t_{\text{cut}}$. This is expressed as $\{p\}_{X+n} > t_{\text{cut}}$ in the above equation ³. The cut

³Here it is assumed that the hard process cross section $\sigma(X)$ is finite everywhere in its phase space.

off scale t_{cut} is called the merging scale. Notice that the constraint functions in eq. (2.18) are different from those in eq. (2.16), since some part of the constraints are already included in the exact tree level cross sections.

Eq. (2.18) summarizes the basic idea of tree level merging algorithms with our notations. Tree level merging algorithms are designed as to improve the DGLAP equation with a help from exact tree level cross sections, or in other words exact tree level matrix elements.

2.2 The CKKW-L algorithm and its extension

In this section, at first the event generation procedure following the improved DGLAP evolution equation in eq. (2.18) by using the CKKW-L merging algorithm [18, 19, 20] is reviewed. Then, we present a new merging algorithm, which differs from the CKKW-L algorithm in the strategy for phase space separation. Our strategy is introduced with the aim of predicting jet angular correlations more accurately. The independence on parton shower starting scale is discussed at the end.

We will examine each term, one by one, in the right hand side (RHS) of eq. (2.18) in the following. In order to generate an event sample according to the probability of the first term

$$\sigma(X) \Pi(t_X, t_{\text{cut}}; \{p\}_X), \quad (2.19)$$

at first an event sample $\{p\}_X$ is generated with the cross section $\sigma(X)$. The next step is to calculate the Sudakov form factor $\Pi(t_X, t_{\text{cut}}; \{p\}_X)$ ⁴. The CKKW-L algorithm uses a parton shower generator to calculate Sudakov form factors. We execute a shower generator on the $\{p\}_X$, setting the shower starting scale to the t_X . If the first evolution scale randomly chosen is higher than the t_{cut} , then we throw away the event sample as a whole and go to the next new event sample. This procedure is equivalent to accepting the event sample with the probability equal to $\Pi(t_X, t_{\text{cut}}; \{p\}_X)$, since it is the no radiation probability during the scale evolution of the $\{p\}_X$ between the scales t_X and t_{cut} . If the first evolution scale is lower than the t_{cut} , this evolution is continued until the cut off scale t_{had} of the shower model and the event sample contributes to the inclusive event samples.

Next, let us look at the second term

$$\sigma(X+1; \{p\}_{X+1} > t_{\text{cut}}) \Pi(t_X, t_1; \{p\}_X) f'(z_1, t_1; \{p\}_X) \Pi(t_1, t_{\text{cut}}; \{p\}_{X+1}). \quad (2.20)$$

At first an event samples $\{p\}_{X+1}$ is generated with the tree level cross section $\sigma(X+1; \{p\}_{X+1} > t_{\text{cut}})$. In order to calculate the first Sudakov form factor $\Pi(t_X, t_1; \{p\}_X)$, we need the intermediate event $\{p\}_X$ and the scale t_1 . In the CKKW-L algorithm, these are obtained from the $\{p\}_{X+1}$ by executing a program which does the exact inverse of the shower generation of a parton shower generator which is used. The program should produce a $\{p\}_X$ and a scale t_1 from the $\{p\}_{X+1}$ as if the parton shower generator had evolved the $\{p\}_X$ and then had generated the $\{p\}_{X+1}$ with the evolution scale t_1 . This backward flow is often called a parton shower history. The construction of the parton shower history completely depends on the shower generator which is used. In Section 2.3, the construction of the PYTHIA8 parton shower history is described. For now let us assume that we successfully construct a shower history of the $\{p\}_{X+1}$ and obtain a $\{p\}_X$ and a scale t_1 . In order to calculate the first Sudakov form factor, the shower generator is executed on the $\{p\}_X$, starting from the scale t_X , as before. If the first evolution scale randomly chosen is higher than the t_1 , we throw away the event sample $\{p\}_{X+1}$ as whole and go to the next new event sample. If not, the constraint function $f'(z_1, t_1; \{p\}_X)$ is calculated and the event sample $\{p\}_{X+1}$ is re-weighted according to it. The calculation of the constraint function is discussed in Section 2.4. If the event sample is still survived after the re-weighting, the second Sudakov form factor $\Pi(t_1, t_{\text{cut}}; \{p\}_{X+1})$ is calculated by executing the shower generator on the $\{p\}_{X+1}$ from the scale t_1 and throwing away the event sample as a whole if the first evolution scale is higher than the t_{cut} . If the first evolution scale is lower than the t_{cut} , this evolution is continued until the cut off scale t_{had} and the event sample contributes to the inclusive event samples.

⁴Strictly speaking, this factor defined in eq. (2.12) is not identical to the Sudakov form factor defined in eq. (2.2) in our notations. However, we call it the Sudakov form factor, too.

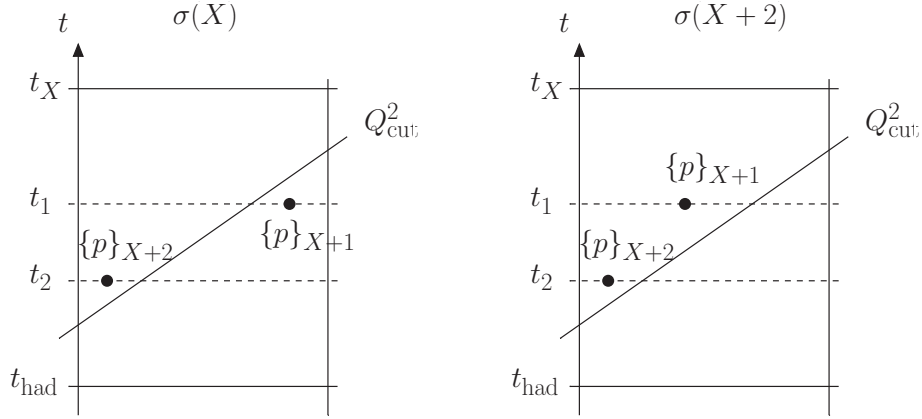


Figure 1: The vertical axis represents the evolution variable t , the horizontal axis represents some variables in the shower model which determine kinematics. The diagonal line indicates the merging scale Q_{cut}^2 . The left and right panels represent the same event $\{p\}_{X+2}$, however their origin can be different. In the CKKW-L algorithm, the contribution in the left panel originates from $\sigma(X)$ and one in the right panel originates from $\sigma(X+2)$. In our merging algorithm, the both originate from $\sigma(X+2)$. See text in Section 2.2 for a detailed explanation.

The event generation following the third and higher terms in the RHS of eq. (2.18) can be performed in the same way. Although merging algorithms can treat tree level cross sections of any number of partons, there are limitations in their calculations. If we decide not to calculate $\sigma(X+3, \{p\}_{X+2} > t_{\text{cut}})$ for instance, what we should do is to remove the last Sudakov form factor $\Pi(t_2, t_{\text{cut}}; \{p\}_{X+2})$ from the third term in the RHS of eq. (2.18) [18]. The third term, which used to be the probability of generating a $\{p\}_{X+2}$ exclusively above the t_{cut} , is now the probability of generating a $\{p\}_{X+2}$ exclusively above a $t_2 (> t_{\text{cut}})$. The evolution equation is closed at the third term, i.e. the dots in eq. (2.18) disappears.

Up to now it has been assumed that the definition of the merging scale t_{cut} is equivalent to that of the shower evolution variable t . It is discussed in refs. [18, 20] that the definition of the merging scale t_{cut} can be arbitrary, as long as it regulates the singularity in tree level matrix elements. We let Q_{cut}^2 denotes a definition of the merging scale. This can be the k_{\perp} definition of the jet clustering algorithm [39] for instance. With an arbitrary definition of the merging scale, the Sudakov form factor in the first term of the RHS of eq. (2.18) is calculated as follows. We execute a shower generator on the $\{p\}_X$, starting from the t_X , as before. After the first evolution, we obtain a $\{p\}_{X+1}$ and calculate the minimum value Q_{min}^2 on the $\{p\}_{X+1}$ by using the definition of the merging scale Q_{cut}^2 . If $Q_{\text{min}}^2 > Q_{\text{cut}}^2$, we throw away the event sample as a whole and go to the next new event sample. We express this as $\{p\}_{X+1} > Q_{\text{cut}}^2$. If not i.e. $Q_{\text{min}}^2 < Q_{\text{cut}}^2$, this evolution is continued until the cut off scale t_{had} and the event sample contributes to the inclusive event samples. We express this as $\{p\}_{X+1} < Q_{\text{cut}}^2$. This procedure is applied to the calculation of the last Sudakov from factor at each term in the RHS of eq. (2.18).

However, it can be easily imagined that, even though $\{p\}_{X+1} < Q_{\text{cut}}^2$ is satisfied after the first evolution of the $\{p\}_X$, the second evolution may generate a $\{p\}_{X+2}$ which is $\{p\}_{X+2} > Q_{\text{cut}}^2$. This is considered as double counting with the contribution from the third term in the RHS of eq. (2.18). The solution to the double counting issue in the CKKW-L algorithm can be explained as follows, by using illustrations in Figure 1⁵. The vertical axis represents the evolution variable t , the horizontal axis represents some kinematic variables in a shower model, such as the energy fraction variable z . One point in the figure corresponds to one phase space point of a $\{p\}_{X+n}$, and it is uniquely determined once a value for the vertical axis and a value for the horizontal axis are chosen. The diagonal line indicates the merging scale Q_{cut}^2 . Note that this line will be perpendicular to the vertical axis, when the definition of the merging scale is equivalent to that of the

⁵This illustration is inspired by refs. [18, 20]

evolution variable i.e. $Q_{\text{cut}}^2 = t_{\text{cut}}$. The left panel shows a contribution of the first term in the RHS of eq. (2.18) after the first and second evolution. This event sample satisfies $\{p\}_{X+1} < Q_{\text{cut}}^2$ after the first evolution, thus it is already accepted. However, after the second evolution, it becomes $\{p\}_{X+2} > Q_{\text{cut}}^2$. The right panel shows a contribution of the third term in the RHS of eq. (2.18) after no evolution. As it is illustrated in the right panel, the CKKW-L algorithm requires intermediate events to satisfy the merging scale cut, i.e. $\{p\}_{X+1} > Q_{\text{cut}}^2$ as well as $\{p\}_{X+2} > Q_{\text{cut}}^2$. The sum of the two contributions in the left and right panels does not lead to double counting, since the $\{p\}_{X+1}$ in the left panel and the $\{p\}_{X+1}$ in the right panel live in the different phase space regions. For the CKKW-L algorithm, eq. (2.18) can be written as follows,

$$\begin{aligned}
& \sigma(X)_{\text{CKKW-L}} \\
&= \sigma(X) \Pi\left(t_X, \{p\}_{X+1} < Q_{\text{cut}}^2; \{p\}_X\right) \\
&+ \sigma\left(X+1; \{p\}_{X+1} > Q_{\text{cut}}^2\right) \Pi(t_X, t_1; \{p\}_X) f'(z_1, t_1; \{p\}_X) \Pi\left(t_1, \{p\}_{X+2} < Q_{\text{cut}}^2; \{p\}_{X+1}\right) \\
&+ \sigma\left(X+2; \{p\}_{X+2} > Q_{\text{cut}}^2, \{p\}_{X+1} > Q_{\text{cut}}^2\right) \Pi(t_X, t_1; \{p\}_X) f'(z_1, t_1; \{p\}_X) \Pi(t_1, t_2; \{p\}_{X+1}) \\
&\quad \times f'(z_2, t_2; \{p\}_{X+1}) \Pi\left(t_2, \{p\}_{X+3} < Q_{\text{cut}}^2; \{p\}_{X+2}\right) \\
&+ \dots
\end{aligned} \tag{2.21}$$

In this study, we use a different method to avoid the double counting issue. When a purpose of merging is to predict angular correlations between two hard jets, it can be one disadvantage of the CKKW-L algorithm that the contribution shown in the left panel of Figure 1 is described by $\sigma(X)$. The $\{p\}_{X+2}$ in the left panel has potential to produce a $X+2$ hard jets event. However, angular correlations between the two jets are not correct, since this event originates from $\sigma(X)$, not $\sigma(X+2)$. Therefore, our algorithm is designed so that the contribution shown in the left panel originates from $\sigma(X+2)$, not $\sigma(X)$. The improved evolution equation in eq. (2.18) for our merging algorithm can be written as follows,

$$\begin{aligned}
& \sigma(X)_{\text{CKKW-L+}} \\
&= \sigma(X) \Pi\left(t_X, \{p\}_{X+1} < Q_{\text{cut}}^2, \{p\}_{X+2} < Q_{\text{cut}}^2, \dots; \{p\}_X\right) \\
&+ \sigma\left(X+1; \{p\}_{X+1} > Q_{\text{cut}}^2\right) \Pi(t_X, t_1; \{p\}_X) f'(z_1, t_1; \{p\}_X) \\
&\quad \times \Pi\left(t_1, \{p\}_{X+2} < Q_{\text{cut}}^2, \{p\}_{X+3} < Q_{\text{cut}}^2, \dots; \{p\}_{X+1}\right) \\
&+ \sigma\left(X+2; \{p\}_{X+2} > Q_{\text{cut}}^2\right) \Pi(t_X, t_1; \{p\}_X) f'(z_1, t_1; \{p\}_X) \Pi(t_1, t_2; \{p\}_{X+1}) \\
&\quad \times f'(z_2, t_2; \{p\}_{X+1}) \Pi\left(t_2, \{p\}_{X+3} < Q_{\text{cut}}^2, \{p\}_{X+4} < Q_{\text{cut}}^2, \dots; \{p\}_{X+2}\right) \\
&+ \dots
\end{aligned} \tag{2.22}$$

Notice that the cut $\{p\}_{X+1} > Q_{\text{cut}}^2$ in the third term of the RHS of eq. (2.21) now disappears and a cut $\{p\}_{X+2} < Q_{\text{cut}}^2$ is newly added in the first term. Merging scale cuts expressed by the dots in the last Sudakov form factor at each term depends on the maximal number of partons predicted by the tree level cross section. In order to make the expression clearer, let us consider the case that we do not calculate $\sigma(X+3)$. The evolution equation is given by

$$\begin{aligned}
& \sigma(X)_{\text{CKKW-L+}}^{N=2} \\
&= \sigma(X) \Pi\left(t_X, \{p\}_{X+1} < Q_{\text{cut}}^2, \{p\}_{X+2} < Q_{\text{cut}}^2; \{p\}_X\right) \\
&+ \sigma\left(X+1; \{p\}_{X+1} > Q_{\text{cut}}^2\right) \Pi(t_X, t_1; \{p\}_X) f'(z_1, t_1; \{p\}_X) \Pi\left(t_1, \{p\}_{X+2} < Q_{\text{cut}}^2; \{p\}_{X+1}\right) \\
&+ \sigma\left(X+2; \{p\}_{X+2} > Q_{\text{cut}}^2\right) \Pi(t_X, t_1; \{p\}_X) f'(z_1, t_1; \{p\}_X) \Pi(t_1, t_2; \{p\}_{X+1}) f'(z_2, t_2; \{p\}_{X+1}), \tag{2.23}
\end{aligned}$$

where N denotes the maximal number of partons predicted by the tree level cross sections. This equation implies that both of the contributions in Figure 1 originate from $\sigma(X+2)$. Since they never originate from

$\sigma(X)$ or $\sigma(X+1)$, there is no double counting. For the sake of completeness, we give the evolution equation for $N=3$,

$$\begin{aligned}
& \sigma(X)_{\text{CKKW-L+}}^{N=3} \\
&= \sigma(X) \Pi(t_X, \{p\}_{X+1} < Q_{\text{cut}}^2, \{p\}_{X+2} < Q_{\text{cut}}^2, \{p\}_{X+3} < Q_{\text{cut}}^2; \{p\}_X) \\
&+ \sigma(X+1; \{p\}_{X+1} > Q_{\text{cut}}^2) \Pi(t_X, t_1; \{p\}_X) f'(z_1, t_1; \{p\}_X) \\
&\quad \times \Pi(t_1, \{p\}_{X+2} < Q_{\text{cut}}^2, \{p\}_{X+3} < Q_{\text{cut}}^2; \{p\}_{X+1}) \\
&+ \sigma(X+2; \{p\}_{X+2} > Q_{\text{cut}}^2) \Pi(t_X, t_1; \{p\}_X) f'(z_1, t_1; \{p\}_X) \Pi(t_1, t_2; \{p\}_{X+1}) \\
&\quad \times f'(z_2, t_2; \{p\}_{X+1}) \Pi(t_2, \{p\}_{X+3} < Q_{\text{cut}}^2; \{p\}_{X+2}) \\
&+ \sigma(X+3; \{p\}_{X+3} > Q_{\text{cut}}^2) \Pi(t_X, t_1; \{p\}_X) f'(z_1, t_1; \{p\}_X) \Pi(t_1, t_2; \{p\}_{X+1}) \\
&\quad \times f'(z_2, t_2; \{p\}_{X+1}) \Pi(t_2, t_3; \{p\}_{X+2}) f'(z_3, t_3; \{p\}_{X+2}). \tag{2.24}
\end{aligned}$$

Notice that a cut $\{p\}_{X+3} < Q_{\text{cut}}^2$ is newly added in the last Sudakov form factor in the first and second terms of the RHS.

In this paper we call the above algorithm given in eq. (2.22) the CKKW-L+ merging algorithm and use it for the event generation. In Section 3.2, the CKKW-L+ algorithm will be numerically compared with the CKKW-L algorithm.

Finally, let us discuss the scale t_X . The DGLAP evolution of a hard process $\sigma(X)$ in eq. (2.16) indicates that the scale t_X is the factorization scale of the hard process. Thus in the merging algorithms the t_X must be determined from the $\{p\}_X$ which is constructed as a parton shower history from a $\{p\}_{X+n}$. Since we obtain the scales t_1, t_2, \dots as shower histories from the event samples generated with tree level matrix elements, it happens that $t_X < t_1, t_2, \dots$. This should be considered as an important prediction of the tree level matrix elements, since this never happens in the DGLAP shower evolution. Even in such a case, the maximal shower evolution scale for the event sample must be set to the t_X . Let us suppose the case that $t_1 > t_X > t_2 > \dots$. The first Sudakov form factor $\Pi(t_X, t_1; \{p\}_X)$ is obviously unity i.e. no veto. A care is needed for the calculation of the second Sudakov form factor. The shower starting scale for the calculation of the second Sudakov form factor $\Pi(t_1, t_2; \{p\}_{X+1})$ or $\Pi(t_1, \{p\}_{X+2} < Q_{\text{cut}}^2, \dots; \{p\}_{X+1})$ must be set to the t_X , not the t_1 . The maximal shower evolution scale t_X is often called parton shower starting scale.

In the shower evolution without merging algorithms as in eq. (2.16), predictions such as a jet p_T distribution can depend strongly on the shower starting scale t_X , since it determines hardness of radiation. It has been confirmed that this dependence is reduced significantly once we use merging algorithms [25]. The reason of the independence can be understood from the improved DGLAP equation in eq. (2.22). Let us rewrite eq. (2.24), which is obtained from eq. (2.22) by setting $N=3$, in the following form

$$\sigma_{\text{inc}}(X) = \sigma_{\text{exc}}(X+0) + \sigma_{\text{exc}}(X+1) + \sigma_{\text{exc}}(X+2) + \sigma_{\text{exc}}(X+3) \tag{2.25}$$

or equivalently

$$1 = \frac{\sigma_{\text{exc}}(X+0)}{\sigma_{\text{inc}}(X)} + \frac{\sigma_{\text{exc}}(X+1)}{\sigma_{\text{inc}}(X)} + \frac{\sigma_{\text{exc}}(X+2)}{\sigma_{\text{inc}}(X)} + \frac{\sigma_{\text{exc}}(X+3)}{\sigma_{\text{inc}}(X)}. \tag{2.26}$$

First of all, the hardness of radiation parametrized by the evolution variable t as t_1, t_2, \dots is determined from the tree level matrix elements, thus is nothing to do with the t_X . Since the first Sudakov form factor at each term such as $\Pi(t_X, t_1; \{p\}_X)$ and $\Pi(t_X, \{p\}_{X+1} < Q_{\text{cut}}^2, \dots; \{p\}_X)$ depends on the t_X , the exclusive cross sections $\sigma_{\text{exc}}(X+i)$ and accordingly the inclusive cross section $\sigma_{\text{inc}}(X)$ are affected by the t_X . However, because the first Sudakov form factor can be considered as an overall factor, the ratio of the exclusive cross section to the inclusive cross section, i.e. $\sigma_{\text{exc}}(X+i)/\sigma_{\text{inc}}(X)$, will be little affected. Stable distribution can be expected from the above two reasons. The dependence on the shower starting scale is numerically evaluated in Section 2.7.

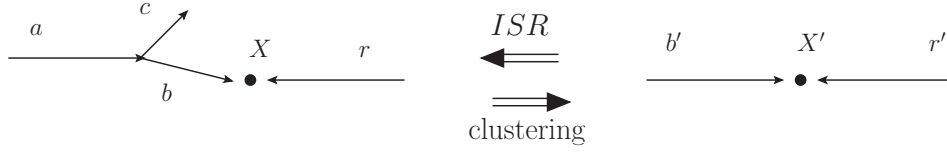


Figure 2: Illustrating that an initial state shower model evolves a process $b'r' \rightarrow X'$ and then generates a process $ar \rightarrow Xc$ (from right to left), and its inverse (from left to right)

2.3 Construction of the PYTHIA8 parton shower history

In Section 2.2, it is explained that the CKKW-L merging algorithms require the construction of the parton shower history and a construction program must do the exact inverse of the shower generation of a parton shower generator which we use. In this section, the construction of the PYTHIA8 parton shower history is described. A parton shower history is constructed by successively clustering two partons into one parton, and one history consists of a set of intermediate events with the corresponding clustering scales which are ordered. For instance, a history of an event sample $\{p\}_{X+n}$ consists of $\{p\}_{X+(n-1)}, \{p\}_{X+(n-2)}, \dots, \{p\}_{X+i}, \dots, \{p\}_{X+1}, \{p\}_X$ with $t_n < t_{n-1} < \dots < t_{i+1} < \dots < t_2 < t_1$. Because the detailed definition of the evolution variable and the kinematics construction are different for initial state radiation (ISR) and final state radiation (FSR) in PYTHIA8, the clustering procedure is also different for a clustering of incoming and outgoing partons and that of two outgoing partons. The former procedure is described in Section 2.3.1 and the latter in Section 2.3.2. In these sections, we use the knowledges and the notations given in the original publications [40, 41, 42] for the shower model in PYTHIA8. Some technical details in our implementation are summarized in Section 2.3.3.

2.3.1 Construction of the ISR history

Let us suppose that an incoming parton a and an outgoing parton c in a process $ar \rightarrow Xc$ are clustered into a new parton b and hence an intermediate process $b'r' \rightarrow X'$ together with a clustering scale $p_{\perp\text{clus}}$ is produced. This is illustrated in Figure 2 (from left to right). The clustering has to proceed as if the initial state shower model in PYTHIA8 had evolved the hard process $b'r' \rightarrow X'$ and then had generated the process $ar \rightarrow Xc$ with the evolution scale $p_{\perp\text{evol}} = p_{\perp\text{clus}}$. The clustering scale $p_{\perp\text{clus}}$ is derived from

$$p_b = p_a - p_c, \quad (2.27a)$$

$$z = \frac{m_{br}^2}{m_{ar}^2} = \frac{(p_b + p_r)^2}{(p_a + p_r)^2}, \quad (2.27b)$$

$$p_{\perp\text{clus}}^2 = -(1-z)(p_b)^2. \quad (2.27c)$$

Here the z can be interpreted as the energy fraction E_b/E_a in the center of mass frame of proton proton collisions.

The new incoming parton b after the clustering is not moving along the z -axis and it is a spacelike particle i.e. $(p_b)^2 < 0$. Thus we need to make the b on-shell (massless) and moving along the z -axis. The X and X' denote all the other particles in the final state, hence $X = X' = t\bar{t} + g$ for instance. The four-momenta of the b' , r' and X' are derived as follows.

1. Read the azimuthal angle ϕ_c of the c .
2. Rotate the c and X in azimuth by $-\phi_c$.
3. Calculate the four-momentum of the b as $p_b = p_a - p_c$.
4. Boost the b , r and X to the $b+r$ rest frame, and then rotate them in polar angle so that the b and r move along the z -axis.

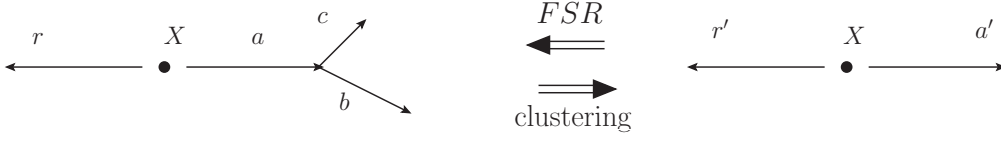


Figure 3: Illustrating that a final state shower model evolves the set of partons a' and r' and then generates partons b , c and r (from right to left), and its inverse (from left to right).

5. Rotate the X in azimuth by $+\phi_c$.

6. Newly obtain the four-momenta of massless incoming partons b and r in the $b + r$ rest frame, i.e. $p_b = (m_{br}/2, 0, 0, m_{br}/2)$ and $p_r = (m_{br}/2, 0, 0, -m_{br}/2)$.

7. Boost the b , r and X along the z -axis so that the r has its original momentum.

With the above algorithm, the non-zero transverse momentum of the parton b is translated into the kinematics of the X . As required, the b' is on mass shell (massless) and it moves along the z -axis. The kinematics of the r does not change i.e. $p_r = p_{r'}$. The above algorithm is carefully tested as follows. We apply the algorithm to an event $ar \rightarrow Xc$ which has been generated by the PYTHIA8 initial state evolution $a \rightarrow bc$ of a hard process event $b'r' \rightarrow X'$, and then we confirm that the algorithm correctly reproduces the event $b'r' \rightarrow X'$.

2.3.2 Construction of the FSR history

Let us consider the case that there are two outgoing partons b and c and one parton r which is either incoming or outgoing. Then let us suppose that the b and c are clustered into a new outgoing parton a and hence the set of partons a' and r' together with a clustering scale $p_{\perp\text{clus}}$ is produced. This is illustrated in Figure 3 (from left to right). The clustering has to proceed as if the final state shower model in PYTHIA8 had evolved the set of the partons a' and r' and then had generated the partons b , c and r with the evolution scale $p_{\perp\text{evol}} = p_{\perp\text{clus}}$. The parton r may not be uniquely determined. In our algorithm, the r is randomly chosen. The new outgoing parton a after the clustering is off mass shell. Thus we need to put the a on mass shell. There are two approaches depending on whether the parton r is outgoing or incoming.

When the r is an outgoing parton, the parton a is put on mass shell by giving the four-momentum of the parton a to the parton r . The four-momentum of the $a + r$ is kept unchanged i.e. $p_a + p_r = p_{a'} + p_{r'}$. The kinematics of all the other partons indicated by X in Figure 3 including the incoming partons will not be affected. The four-momenta of the partons a' and r' are derived as follows.

1. Boost the a and r to the $a + r$ rest frame, $p_{a,0}$ and $p_{r,0}$.

2. In the $a + r$ rest frame, calculate the energies and the absolute values of the momenta of the a and r which are put on mass shell with the on-shell masses m_a and m_r ,

$$m_{ar}^2 = (p_{a,0} + p_{r,0})^2, \quad (2.28a)$$

$$E_{a,\text{new}} = \frac{m_{ar}^2 + m_a^2 - m_r^2}{2m_{ar}}, \quad (2.28b)$$

$$E_{r,\text{new}} = \frac{m_{ar}^2 - m_a^2 + m_r^2}{2m_{ar}}, \quad (2.28c)$$

$$|\vec{p}_{a,\text{new}}| = |\vec{p}_{r,\text{new}}| = \sqrt{E_{a,\text{new}}^2 - m_a^2}. \quad (2.28d)$$

3. Modify the magnitudes of the momenta of the a and r to the $|\vec{p}_{a,\text{new}}|$ and $|\vec{p}_{r,\text{new}}|$, respectively, while the directions of the momenta are kept unchanged,

$$p_a^i = \frac{p_{a,0}^i}{|\vec{p}_{a,0}|} |\vec{p}_{a,\text{new}}|, \quad p_r^i = \frac{p_{r,0}^i}{|\vec{p}_{r,0}|} |\vec{p}_{r,\text{new}}|. \quad (i = 1, 2, 3) \quad (2.29)$$

4. Boost these back to the original $a + r$ frame.

When the r is an incoming parton, the parton a is put on mass shell by reducing the four-momenta of the a and r , while the four-momentum $p_r - p_a$ is kept unchanged. Thus, the four-momentum $p_a + p_r$ will not be conserved in this case. The four-momenta of the partons a' and r' are derived from

$$\alpha = \frac{p_r^3}{E_r} = 1 \text{ or } -1, \quad (2.30a)$$

$$p_{r'}^\mu = \left(E_r - \frac{(p_a)^2 - m_a^2}{2(E_a - \alpha p_a^3)}, 0, 0, \alpha \left(E_r - \frac{(p_a)^2 - m_a^2}{2(E_a - \alpha p_a^3)} \right) \right), \quad (2.30b)$$

$$p_{a'}^\mu = (E_a + E_{r'} - E_r, p_a^1, p_a^2, p_a^3 + p_{r'}^3 - p_r^3). \quad (2.30c)$$

For either case, the clustering scale $p_{\perp \text{clus}}$ is derived from,

$$p_a = p_b + p_c, \quad (2.31a)$$

$$p_0 = p_{a'} + p_{r'}, \quad (2.31b)$$

$$z = \frac{p_0 \cdot p_b}{p_0 \cdot p_a} \left(1 - \frac{p_0 \cdot p_c}{p_0 \cdot p_b} \frac{m_a^2}{2p_0 \cdot p_a + m_r^2 - m_a^2 - p_0^2} \right), \quad (2.31c)$$

$$p_{\perp \text{clus}}^2 = z(1 - z)((p_a)^2 - m_a^2), \quad (2.31d)$$

where m_a and m_r are the on-shell masses of the partons a and r , respectively. Notice that the p_0 is constructed from the four-momenta of the partons a' and r' , which are obtained from the above algorithm. The mass effect is taken into account. This is particularly relevant in the clustering which includes a top quark. Our algorithm is carefully tested as follows. We apply the algorithm to a process which has been generated by the PYTHIA8 final state evolution of a hard process, and then we confirm that the algorithm correctly reproduces the hard process.

2.3.3 Some technical details

In this section, we write some technical details in our implementation of the shower history construction:

- The clustering $2 \rightarrow 1$ must respect the QCD $1 \rightarrow 2$ vertices.
- When there are more than one candidates for a clustering pair at a clustering step, the one which has the lowest clustering scale is always chosen ⁶.
- Sequential clustering scales are required to be ordered, that is, a clustering scale at a clustering step is required to be higher than the scale at the previous clustering step. However, it is not required that $t_X > t_1, t_2, \dots$.
- If the hard process $\{p\}_X$ cannot be obtained at the end of sequential clusterings, we take the following approach. Let us assume the case that the $\{p\}_X$ is not obtained from a $\{p\}_{X+n}$ after sequential n times clusterings. The program for the shower history construction is executed again on the $\{p\}_{X+n}$, but this time a clustering pair whose clustering scale is not the lowest but the second lowest is chosen at the first clustering step. If the $\{p\}_X$ is not obtained yet, the program is executed again on the $\{p\}_{X+n}$ and a clustering pair whose clustering scale is the third lowest is chosen at the first clustering step, and so on. If this approach still does not help, the shower history construction for the event sample is abandoned.

⁶A more sophisticated approach is proposed in ref. [20].

2.4 Weight functions

In eq. (2.10) of Section 2.1, we have introduced a function $f(z, t; \{p\}_{X+n})$ which constraints the DGLAP shower evolution of the $\{p\}_{X+n}$ at the evolution scale t and the energy fraction z . In the improved DGLAP equation, some part of the constraint is already included in the tree level cross sections. This has been implied by using $f'(z, t; \{p\}_{X+n})$ instead of $f(z, t; \{p\}_{X+n})$ in eqs. (2.18) or (2.22). To make it clear, let us write down the second term in the right hand side (RHS) of the DGLAP evolution of the hard process $\sigma(X)$ in eq. (2.16), which gives the integrated probability of exclusively generating one radiation during the evolution between t_X and t_Λ ,

$$\int_0^1 dx_1 \int_0^1 dx_2 g(x_1, t_X) g(x_2, t_X) \hat{\sigma}(gg \rightarrow X; \hat{s} = x_1 x_2 s) \int_{t_\Lambda}^{t_X} \frac{dt_1}{t_1} \int_0^1 d\hat{p}(z_1) f(z_1, t_1; \{p\}_X), \quad (2.32)$$

where now the hard process is explicitly written and the Sudakov form factors are omitted. This term is improved with a help from the tree level cross section

$$\int_0^1 dx_1 \int_0^1 dx_2 g(x_1, t_1) g(x_2, t_1) \hat{\sigma}(gg \rightarrow X + g; \hat{s} = x_1 x_2 s) f'(z_1, t_1; \{p\}_X), \quad (2.33)$$

where the merging scale cut $\{p\}_{X+g} > Q_{\text{cut}}^2$ is implicit. Furthermore, we write this as

$$\int_0^1 dx_1 \int_0^1 dx_2 g(x_1, t_\Lambda) g(x_2, t_\Lambda) \hat{\sigma}(gg \rightarrow X + g; \hat{s} = x_1 x_2 s) f'(z_1, t_1; \{p\}_X) f'(t_\Lambda; \{p\}_{X+g}). \quad (2.34)$$

A fixed scale t_Λ is now used in the PDFs and this change is included in the function $f'(t_\Lambda; \{p\}_{X+g})$. We call $f'(z, t; \{p\}_{X+n})$ a weight function hereafter, since it is used to re-weight an event sample in the merging algorithms, see the discussion below eq. (2.20) for the detail. The first task in this section is to derive the weight function $f'(z_1, t_1; \{p\}_X)$ explicitly.

There are three radiation patterns in the PYTHIA8 parton shower, namely final state radiation with an outgoing recoiling parton, initial state radiation with an incoming recoiling parton and final state radiation with an incoming recoiling parton. This is also discussed in Section 2.3. For the first case i.e. final state radiation with an outgoing recoiling parton, the kinematics of the incoming partons will not be changed. Hence eq. (2.32) is evaluated as

$$\begin{aligned} & \int_0^1 dx_1 \int_0^1 dx_2 g(x_1, t_X) g(x_2, t_X) \hat{\sigma}(gg \rightarrow X; \hat{s} = x_1 x_2 s) \int_{t_\Lambda}^{t_X} \frac{dt_1}{t_1} \int_0^1 d\hat{p}(z_1) \\ &= \int_0^1 dx_1 \int_0^1 dx_2 g(x_1, t_X) g(x_2, t_X) \hat{\sigma}(gg \rightarrow X + g; \hat{s} = x_1 x_2 s) \\ &= \int_0^1 dx_1 \int_0^1 dx_2 g(x_1, t_\Lambda) g(x_2, t_\Lambda) \hat{\sigma}(gg \rightarrow X + g; \hat{s} = x_1 x_2 s) \frac{g(x_1, t_X)}{g(x_1, t_\Lambda)} \frac{g(x_2, t_X)}{g(x_2, t_\Lambda)}, \end{aligned} \quad (2.35)$$

thus we obtain

$$f'(z_1, t_1; \{p\}_X) f'(t_\Lambda; \{p\}_{X+g}) = \frac{g(x_1, t_X)}{g(x_1, t_\Lambda)} \frac{g(x_2, t_X)}{g(x_2, t_\Lambda)}, \quad (2.36)$$

or equivalently

$$f'(z_1, t_1; \{p\}_X) f'(t_\Lambda; \{p\}_{X+g}) = \frac{g(x_1, t_1)}{g(x_1, t_\Lambda)} \frac{g(x_1, t_X)}{g(x_1, t_1)} \frac{g(x_2, t_X)}{g(x_2, t_1)} \frac{g(x_2, t_1)}{g(x_2, t_\Lambda)}. \quad (2.37)$$

For the second case i.e. initial state radiation with an incoming recoiling parton, the kinematics of a radiating incoming parton will be changed. Thus the weight function should include the parton distribution

functions (PDFs). The PDF factor is explicitly given in eq. (2.8). When we assume that the incoming parton which has the energy fraction x_1 is the radiating parton, eq. (2.32) is evaluated as

$$\begin{aligned}
& \int_0^1 dx_1 \int_0^1 dx_2 g(x_1, t_X) g(x_2, t_X) \hat{\sigma}(gg \rightarrow X; \hat{s} = x_1 x_2 s) \int_{t_\Lambda}^{t_X} \frac{dt_1}{t_1} \int_0^1 d\hat{p}(z_1) \frac{g(x_1/z_1, t_1)}{g(x_1, t_1)} \\
&= \int_0^1 dx_1 \int_0^1 dx_2 g(x_1, t_X) g(x_2, t_X) \hat{\sigma}(gg \rightarrow X + g; \hat{s} = (x_1/z_1) x_2 s) \frac{1}{z_1} \frac{g(x_1/z_1, t_1)}{g(x_1, t_1)} \\
&= \int_0^1 dw_1 \int_0^1 dx_2 g(z_1 w_1, t_X) g(x_2, t_X) \hat{\sigma}(gg \rightarrow X + g; \hat{s} = w_1 x_2 s) \frac{g(w_1, t_1)}{g(z_1 w_1, t_1)} \\
&= \int_0^1 dw_1 \int_0^1 dx_2 g(w_1, t_1) g(x_2, t_1) \hat{\sigma}(gg \rightarrow X + g; \hat{s} = w_1 x_2 s) \frac{g(z_1 w_1, t_X)}{g(z_1 w_1, t_1)} \frac{g(x_2, t_X)}{g(x_2, t_1)} \\
&= \int_0^1 dw_1 \int_0^1 dx_2 g(w_1, t_\Lambda) g(x_2, t_\Lambda) \hat{\sigma}(gg \rightarrow X + g; \hat{s} = w_1 x_2 s) \frac{g(w_1, t_1)}{g(w_1, t_\Lambda)} \frac{g(z_1 w_1, t_X)}{g(z_1 w_1, t_1)} \frac{g(x_2, t_X)}{g(x_2, t_1)} \frac{g(x_2, t_1)}{g(x_2, t_\Lambda)},
\end{aligned} \tag{2.38}$$

thus we obtain

$$f'(z_1, t_1; \{p\}_X) f'(t_\Lambda; \{p\}_{X+g}) = \frac{g(w_1, t_1)}{g(w_1, t_\Lambda)} \frac{g(z_1 w_1, t_X)}{g(z_1 w_1, t_1)} \frac{g(x_2, t_X)}{g(x_2, t_1)} \frac{g(x_2, t_1)}{g(x_2, t_\Lambda)}. \tag{2.39}$$

Note that $1/z_1$ at the second line of eq. (2.38) comes from $d\hat{p}(z_1)$, see eq. (2.13).

For the third case i.e. final state radiation with an incoming recoiling parton, the kinematics of the incoming recoiling parton will be changed. Thus the weight function includes the PDFs [41]. When we assume that the incoming parton which has the energy fraction x_1 is the recoiling parton, eq. (2.32) is evaluated as, by letting w_1 denotes the energy fraction of the recoiling parton after the radiation,

$$\begin{aligned}
& \int_0^1 dx_1 \int_0^1 dx_2 g(x_1, t_X) g(x_2, t_X) \hat{\sigma}(gg \rightarrow X; \hat{s} = x_1 x_2 s) \int_{t_\Lambda}^{t_X} \frac{dt_1}{t_1} \int_0^1 d\hat{p}(z_1) \frac{w_1}{x_1} \frac{g(w_1, t_1)}{g(x_1, t_1)} \\
&= \int_0^1 dw_1 \int_0^1 dx_2 g(x_1, t_X) g(x_2, t_X) \hat{\sigma}(gg \rightarrow X; \hat{s} = x_1 x_2 s) \int_{t_\Lambda}^{t_X} \frac{dt_1}{t_1} \int_0^1 d\hat{p}(z_1) \frac{g(w_1, t_1)}{g(x_1, t_1)} \\
&= \int_0^1 dw_1 \int_0^1 dx_2 g(w_1, t_1) g(x_2, t_X) \hat{\sigma}(gg \rightarrow X; \hat{s} = x_1 x_2 s) \int_{t_\Lambda}^{t_X} \frac{dt_1}{t_1} \int_0^1 d\hat{p}(z_1) \frac{g(x_1, t_X)}{g(x_1, t_1)} \\
&= \int_0^1 dw_1 \int_0^1 dx_2 g(w_1, t_1) g(x_2, t_X) \hat{\sigma}(gg \rightarrow X + g; \hat{s} = w_1 x_2 s) \frac{g(x_1, t_X)}{g(x_1, t_1)} \\
&= \int_0^1 dw_1 \int_0^1 dx_2 g(w_1, t_\Lambda) g(x_2, t_\Lambda) \hat{\sigma}(gg \rightarrow X + g; \hat{s} = w_1 x_2 s) \frac{g(w_1, t_1)}{g(w_1, t_\Lambda)} \frac{g(x_1, t_X)}{g(x_1, t_1)} \frac{g(x_2, t_X)}{g(x_2, t_1)} \frac{g(x_2, t_1)}{g(x_2, t_\Lambda)},
\end{aligned} \tag{2.40}$$

thus we obtain

$$f'(z_1, t_1; \{p\}_X) f'(t_\Lambda; \{p\}_{X+g}) = \frac{g(w_1, t_1)}{g(w_1, t_\Lambda)} \frac{g(x_1, t_X)}{g(x_1, t_1)} \frac{g(x_2, t_X)}{g(x_2, t_1)} \frac{g(x_2, t_1)}{g(x_2, t_\Lambda)}. \tag{2.41}$$

By looking at the weight functions of the three radiation patterns in eqs. (2.37), (2.39) and (2.41), the general expression can be found as follows. Let us suppose an event sample $\{p\}_{X+n}$, which gives a shower history $\{p\}_{X+(n-1)}, \{p\}_{X+(n-2)}, \dots, \{p\}_{X+i}, \dots, \{p\}_{X+1}, \{p\}_X$ with $t_n < t_{n-1} < \dots < t_{i+1} < \dots < t_2 < t_1$. Let us also define the energy fractions and the parton types of the incoming partons in the $\{p\}_{X+i}$ by $x_1^{(i)}$, $x_2^{(i)}$ and $f_1^{(i)}$, $f_2^{(i)}$, respectively. The weight function for the $\{p\}_{X+i}$ is given by [20]

$$f'(z_{i+1}, t_{i+1}; \{p\}_{X+i}) = \frac{\alpha_s(t_{i+1})}{\alpha_s(t_\Lambda)} \frac{f_1^{(i)}(x_1^{(i)}, t_i)}{f_1^{(i)}(x_1^{(i)}, t_{i+1})} \frac{f_2^{(i)}(x_2^{(i)}, t_i)}{f_2^{(i)}(x_2^{(i)}, t_{i+1})}, \tag{2.42}$$



Figure 4: Schematic pictures showing that two different shower histories are constructed for an event $gg \rightarrow t\bar{t}g$. In the left picture, the clustering $t + g \rightarrow t$, which gives the lowest clustering scale t_1 of all the candidates for a clustering pair in the event, is performed. The right picture shows the case that the clustering $t + g \rightarrow t$ is not allowed in the program. A different clustering which gives a higher scale t'_1 than the t_1 will be performed instead for the same event.

from which the total weight function for the $\{p\}_{X+n}$ is

$$\prod_{i=0}^n f'(z_{i+1}, t_{i+1}; \{p\}_{X+i}), \quad (2.43)$$

where $t_0 = t_X$ and $t_{n+1} = t_\Lambda$. It is assumed that the scale t_Λ is used as the scales of the PDFs and those of the strong couplings for generating the $\{p\}_{X+n}$. A factor consisting of the strong couplings is present in the function, because the evolution scale t is used as the scale in the strong coupling in the PYTHIA8 parton shower evolution and thus this strategy should be also used in the improved evolution equations [14]. We take into account the fact that The PYTHIA8 shower model uses the different strong couplings for initial state radiation and final state radiation. When a radiation is classified as the initial state radiation (the final state radiation) with a scale t_{i+1} by the shower history construction, $\alpha_s(m_z) = 1.37$ (1.383) is used for the factor in eq. (2.42). It can be easily confirmed that the results in eqs. (2.37), (2.39) and (2.41) are derived from eq. (2.43) by setting $n = 1$.

2.5 The merging algorithm with top decays

The decay of the top quark is characterized by its decay width $\Gamma(t \rightarrow bW^+) \sim 1.5$ GeV. The large scale discrepancy between the decay width and the production scale of a top quark ($\sim m_t$) indicates that QCD radiation off a top quark takes place faster than the decay of the top quark. Although gluon radiation off a top quark is suppressed by its large mass, it can be important for more accurate predictions. The parton shower generator in PYTHIA8 models gluon radiation off heavy particles including the top quark [42]. Therefore, the construction of the PYTHIA8 parton shower history also has to take into account the clustering which includes a top quark i.e. $t + g \rightarrow t$.

Let us first examine the case that the clustering $t + g \rightarrow t$ is not implemented in the program for the shower history construction. Figure 4 presents schematic pictures showing that two different shower histories are constructed for an event $gg \rightarrow t\bar{t}g$. In the left picture, the clustering $t + g \rightarrow t$ which gives the lowest clustering scale t_1 of all the candidates for a clustering pair in the event is performed. If the clustering $t + g \rightarrow t$ is not implemented in the program, a different clustering which gives a higher scale t'_1 than the t_1 will be performed for the same event, as illustrated in the right picture. This difference in the parton shower history affects the calculation of the Sudakov form factors in the evolution equation in eq. (2.22) and thus can induce some problems such as the unnecessary dependence on the merging scale.

The clustering $t + g \rightarrow t$ in the construction of the PYTHIA8 parton shower history is necessary, no matter whether top quarks are decayed in simulation, since the decays occur after gluon radiation off the top quarks as discussed above. The simplest approach to simulate the $t\bar{t}$ production including a $t\bar{t}$ decay may be as follows. At first, the event samples of the $t\bar{t}$ production process are generated with the merging algorithms i.e. $X = t\bar{t}$ in eq. (2.22), by assuming that the $t\bar{t}$ is stable. Then, the top and antitop quarks are decayed independently following the differential partial decay width, $d\Gamma(t \rightarrow l^+\bar{\nu}b)$ or $d\Gamma(t \rightarrow q\bar{q}'b)$. QCD radiation off the decay products will be generated at the end. In this simple approach, correlations between

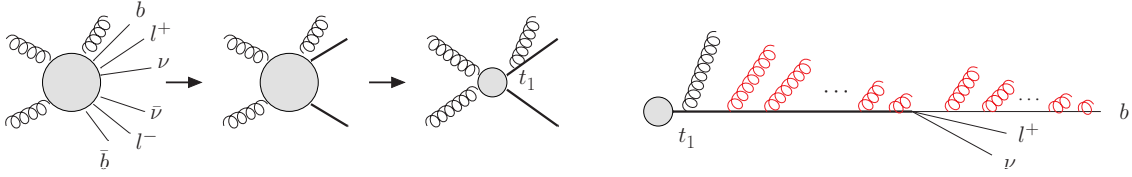


Figure 5: *left*: An event sample $\{p\}_{bl+\nu\bar{b}l-\bar{\nu}+g}$ reconstructs a $\{p\}_{t\bar{t}+g}$ as an intermediate process, at the first allow. A shower history of the $\{p\}_{t\bar{t}+g}$ is constructed, at the second allow. *right*: A part of the shower evolution of the $\{p\}_{bl+\nu\bar{b}l-\bar{\nu}+g}$. The black lines represent the particles generated according to the tree level cross section. The red lines represent particles generated with the shower evolution.

the decay products of the top quark and those of the antitop quark are not produced correctly, and the off shell effects of the top and antitop quarks are also absent. One solution to these issues is to generate the event samples which include a $t\bar{t}$ decay as a part of the hard process, according to the exact tree level matrix elements.

In our study, the event generation of the $t\bar{t}$ production including a $t\bar{t}$ decay is performed as follows. Let us consider an event sample $\{p\}_{bl+\nu\bar{b}l-\bar{\nu}+g}$, which originates from a $\{p\}_{t\bar{t}+g}$ and is generated with the tree level matrix elements. At first, the $\{p\}_{t\bar{t}+g}$ is reconstructed as an intermediate process from the $\{p\}_{bl+\nu\bar{b}l-\bar{\nu}+g}$. The top and antitop quarks are not necessarily on-shell. Next, the program for the shower history construction is executed on the $\{p\}_{t\bar{t}+g}$, and hence a $\{p\}_{t\bar{t}}$ and a clustering scale t_1 are obtained as a shower history. The sequence of the procedures is illustrated in the left picture of Figure 5. The first Sudakov form factor in the evolution equation in eq. (2.22) will be calculated based on the $\{p\}_{t\bar{t}}$ and the t_1 , i.e. $\Pi(t_X, t_1; \{p\}_{t\bar{t}})$. The calculation of the second Sudakov form factor $\Pi(t_1, \{p\}_{t\bar{t}+2} < Q_{\text{cut}}^2, \dots; \{p\}_{t\bar{t}+g})$ looks a little tricky, because the shower generator must be executed on the $\{p\}_{t\bar{t}+g}$ for the calculation of the second Sudakov form factor, while it must be executed on the $\{p\}_{bl+\nu\bar{b}l-\bar{\nu}+g}$ for the event generation. In fact, when the top and antitop quarks are present as intermediate particles in a Les Houches event file of the $\{p\}_{bl+\nu\bar{b}l-\bar{\nu}+g}$, the PYTHIA8 parton shower program starts the shower evolution of the $\{p\}_{t\bar{t}+g}$ at first. This implementation makes the calculation of $\Pi(t_1, \{p\}_{t\bar{t}+2} < Q_{\text{cut}}^2, \dots; \{p\}_{t\bar{t}+g})$ possible. Once the shower evolution of the $\{p\}_{t\bar{t}+g}$ is completed, the kinematic change of the $t\bar{t}$ due to the evolution is reflected into the kinematics of its decay products i.e. $bl+\nu\bar{b}l-\bar{\nu}$. Finally, the shower evolution of the decay products is performed in such a way that it does not change the invariant mass of the top quark and that of the antitop quark [40]. The right picture of Figure 5 schematically shows a part of the shower evolution of the $\{p\}_{bl+\nu\bar{b}l-\bar{\nu}+g}$. The black lines represent the particles generated according to the tree level cross section. The red lines represent particles generated with the shower evolution. Note that the first radiation off the bottom quark is also corrected internally by using the differential decay width $d\Gamma(t \rightarrow bW^+g)$ [42].

2.6 Event generation

In this section, we describe a procedure for the event generation of the top quark pair production in detail, by combining the knowledges given in the above sections.

1. Generate the event samples for the $t\bar{t}+0, 1, \dots, N$ partons production processes at proton proton (pp) collisions according to the tree level cross sections, i.e. $\{p\}_{t\bar{t}}, \{p\}_{t\bar{t}+1}, \dots, \{p\}_{t\bar{t}+N}$. When a decay of the $t\bar{t}$ is to be simulated, the event samples are generated according to the tree level cross sections including the decay of the $t\bar{t}$, i.e. $\{p\}_{bl+\nu\bar{b}l-\bar{\nu}}, \{p\}_{bl+\nu\bar{b}l-\bar{\nu}+1}, \dots, \{p\}_{bl+\nu\bar{b}l-\bar{\nu}+N}$. In this paper, only the dilepton decay is studied. We let N denotes the maximal number of partons provided by the tree level cross section. We use MadGraph5_aMC@NLO[43] version 5.2.2.1 for this purpose. The merging scale

Q_{cut} is defined by the longitudinal-boost invariant k_{\perp} variable [39]

$$k_{\perp iB} = p_{Ti}, \quad (2.44a)$$

$$k_{\perp ij} = \min(p_{Ti}, p_{Tj}) \sqrt{(y_i - y_j)^2 + (\phi_i - \phi_j)^2} / R, \quad (2.44b)$$

where p_{Ti} , y_i and ϕ_i are the transverse momentum with respect to the beam, rapidity and azimuthal angle of outgoing particle i . R is the radius parameter and $R = 1$ is used if not otherwise specified. The merging scale cut is imposed only on the light partons, and no cut is imposed on the $t\bar{t}$ and its decay products. A fixed value t_{Λ} is used for the scales in the strong couplings and in the parton distribution functions (PDFs). The PDF set CTEQ6L1 [44] is used. The center of mass energy for the pp collisions is 14 TeV, except in the case that the simulation is compared with the data at the 7 TeV.

2. Select an event sample for the $t\bar{t} + n$ partons process, i.e. $\{p\}_{t\bar{t}+n}$, or $\{p\}_{bl+\nu\bar{b}l-\bar{\nu}+n}$ when the $t\bar{t}$ is decayed, with the probability proportional to its integrated tree level cross section obtained in step 1,

$$P_n = \frac{\sigma(pp \rightarrow t\bar{t} + n)}{\sum_{i=0}^N \sigma(pp \rightarrow t\bar{t} + i)}. \quad (2.45)$$

3. Construct a parton shower history of the $\{p\}_{t\bar{t}+n}$ by following the procedure described in Section 2.3. The history consists of intermediate events $\{p\}_{t\bar{t}+(n-1)}, \{p\}_{t\bar{t}+(n-2)}, \dots, \{p\}_{t\bar{t}+i}, \dots, \{p\}_{t\bar{t}+1}, \{p\}_{t\bar{t}}$ with the scales $t_n < t_{n-1} < \dots < t_{i+1} < \dots < t_2 < t_1$. When the $t\bar{t}$ is decayed, a $\{p\}_{t\bar{t}+n}$ is reconstructed at first from the $\{p\}_{bl+\nu\bar{b}l-\bar{\nu}+n}$ as described in Section 2.5.
4. When the CKKW-L algorithm is used, the merging scale cut is imposed on the intermediate events, too, as indicated in eq. (2.21). Thus, the event sample is vetoed as a whole, unless $\{p\}_{t\bar{t}+i} > Q_{\text{cut}}^2$ for $i = 1, 2, \dots, n-2, n-1$. This is not the case in the CKKW-L+ algorithm.
5. Calculate the weight function for the $\{p\}_{t\bar{t}+n}$ by using eqs. (2.42) and (2.43). The scale t_X must be determined from the intermediate event $\{p\}_{t\bar{t}}$ and we use

$$t_X = E_T(t) \times E_T(\bar{t}), \quad (2.46)$$

where $E_T^2 = m^2 + p_T^2$. Note that when the $t\bar{t}$ is on-shell, $E_T(t) = E_T(\bar{t})$ thus $t_X = E_T^2(t)$. This is not the case when the decay of the $t\bar{t}$ is also generated in step 1, since the reconstructed $t\bar{t}$ is not necessarily on-shell. This scale t_X is also used as the renormalization scales of the strong couplings α_s^2 for the hard process, that is,

$$\frac{\alpha_s^2(t_X)}{\alpha_s^2(t_{\Lambda})} \quad (2.47)$$

should be added in the weight function. We use $\alpha_s(m_z) = 0.13$ in the above factor. When the construction of the shower history was abandoned in step 3, it is not possible to calculate the weight function in the given way. In such a case, the weight function is given by, instead of eqs. (2.42) and (2.43),

$$\frac{\alpha_s^2(t_X)}{\alpha_s^2(t_{\Lambda})} \frac{\alpha_s(p_T^2(1))\alpha_s(p_T^2(2)) \cdots \alpha_s(p_T^2(n))}{\alpha_s^n(t_{\Lambda})} \frac{f_1^{(n)}(x_1^{(n)}, t_X)}{f_1^{(n)}(x_1^{(n)}, t_{\Lambda})} \frac{f_2^{(n)}(x_2^{(n)}, t_X)}{f_2^{(n)}(x_2^{(n)}, t_{\Lambda})}, \quad (2.48)$$

where $p_T(i)$ is the transverse momentum of a parton i in the $\{p\}_{t\bar{t}+n}$. The scale t_X is defined by eq. (2.46) and now is determined from the $\{p\}_{t\bar{t}+n}$. We use $\alpha_s(m_z) = 0.13$ for all the strong couplings in eq. (2.48). Once the weight function is obtained, the $\{p\}_{t\bar{t}+n}$ is re-weighted with the function. However, the weight function is not bounded above by unity. Therefore, the upper bound of the weight function must be found at first by calculating the weight function for a large number of $\{p\}_{t\bar{t}+n}$. The integrated cross section obtained in step 1 has to be multiplied by the obtained upper bound of the weight function, of course.

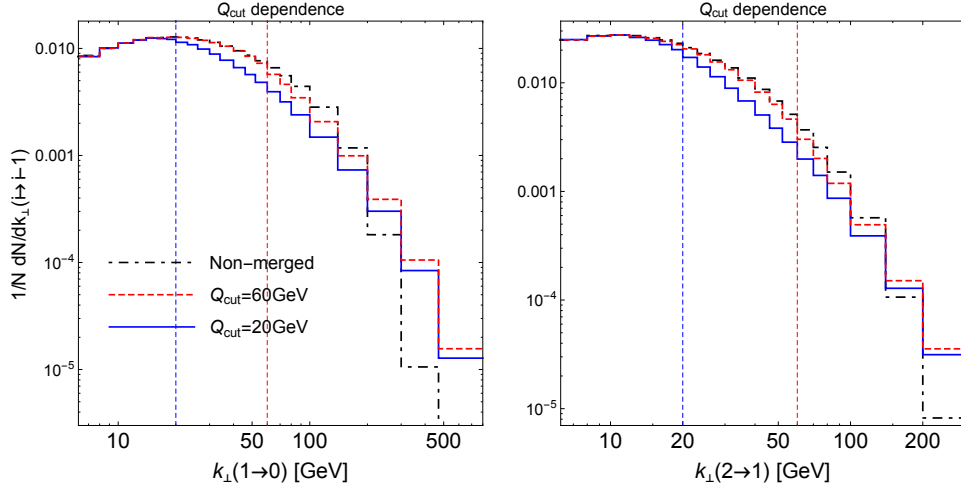


Figure 6: The differential jet rates for $1 \rightarrow 0$ (left) and $2 \rightarrow 1$ (right) jets. The merging scale is set to $Q_{\text{cut}} = 20$ GeV for the blue solid curve and to $Q_{\text{cut}} = 60$ GeV for the red dashed curve. The black broken curve represents the result without merging algorithms

6. Calculate the Sudakov form factor(s) by following the procedure described in Section 2.2. When the construction of the shower history was abandoned in step 3, all the Sudakov form factors are set to unity. We use the parton shower model [40, 41, 42] in PYTHIA8 [45, 46] version 8186. The default tune of the version 8186, tune 4C [41], is basically used, while some functions are turned off. To simplify the analysis, the hadronization after the shower evolution and the multiple interaction are turned off. The rapidity ordering in the initial state radiation is turned off as suggested in ref. [20]. All functions inducing azimuthal asymmetry are turned off, since azimuthal angle information of hard partons is provided by exact tree level matrix elements in our simulation.
7. Once the event sample is accepted and thus the shower evolution is performed until the shower cutoff scale in step 6, all visible particles in the final state within a rapidity range $|y| < 5.0$ including charged leptons are clustered to construct inclusive jets according to the anti- k_T algorithm [47]. The $t\bar{t}$ will not be included, if it is not decayed. The radius parameter is $R = 0.4$ if not otherwise specified. We use Fastjet [48] version 3.1.0 for this purpose. The rapidity and p_T cuts on jets will be specified in the studies of Section 3.
8. Repeat the above procedures from step 2 to step 7 until a large number of the accepted event samples are accumulated.

2.7 Tests of our implementation

In this section, our implementation of the CKKW-L+ merging algorithm is tested. At first, the dependence of differential jet rates on the merging scale cutoff Q_{cut} and that on the parton shower starting scale are studied. A comparison with experimental data is also presented.

Differential jet rates are calculated by using the longitudinal-boost invariant k_{\perp} definition in eq. (2.44) with the radius parameter $R = 1$. In Figure 6 the differential jet rates for $1 \rightarrow 0$ (left) and $2 \rightarrow 1$ (right) jets are plotted. The maximal number of partons N predicted by the tree level cross section is $N = 3$, i.e. the event samples are generated exactly according to eq. (2.24). A vertical dashed line indicates the merging scale Q_{cut} . The merging scale is set to $Q_{\text{cut}} = 20$ GeV for the blue solid curve and to $Q_{\text{cut}} = 60$ GeV for the red dashed curve. The black broken curve represents the result without merging algorithms i.e. purely the

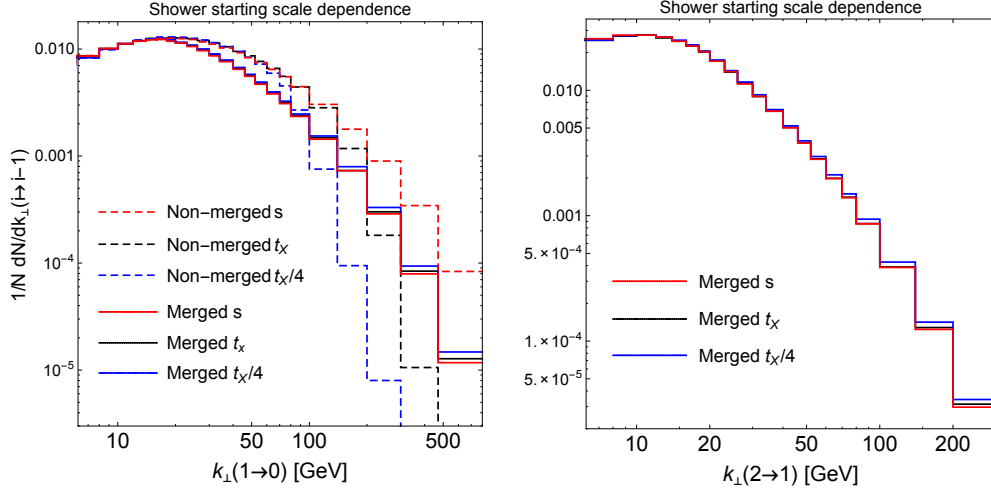


Figure 7: The differential jet rates for $1 \rightarrow 0$ (left) and $2 \rightarrow 1$ (right) jets. The dependence on the parton shower starting scale is studied. The solid lines represent the results obtained with the merging algorithm and the dashed lines represent the results obtained without merging algorithms. Three different scale choices are considered for the shower starting scale, namely $s = (14)^2 \text{ TeV}^2$, t_X and $t_X/4$.

shower prediction. The three curves are set equal at the bin between 10 and 12 GeV for comparison.

The obtained results are distributed smoothly around the merging scales. A clear difference is, however, observed by varying the merging scale from 20 GeV to 60 GeV. The reason for this is that the results obtained with the merging algorithm already deviate from the parton shower prediction in the soft and/or collinear region. By comparing the blue solid curve and the black broken curve, it is clear that the result obtained with the merging algorithm starts to deviate from the shower prediction at around the merging scale (20 GeV). The same can be observed for the red dashed curve. This behavior is quite natural, considering the idea of merging algorithms, namely parton showers are populated below the merging scale and tree level matrix elements are populated above the merging scale. The observation of the clear difference does not imply a fault in our implementation of the CKKW-L+ merging algorithm, but suggests that we should choose a smaller value for the merging scale for the $t\bar{t}$ pair production when the current shower model is used.

As the second test, the dependence on the parton shower starting scale is studied. The parton shower starting scale has been explained at the end of Section 2.2. Although the most natural choice for the parton shower starting scale is the scale t_X , we have argued that predictions such as a jet p_T distribution should be insensitive to the parton shower starting scale. We have also mentioned that the inclusive cross section can be sensitive to it. We confirm these statements in the following. We consider three different scale choices for the parton shower starting scale, namely $s = (14)^2 \text{ TeV}^2$, t_X and $t_X/4$. We use $N = 3$ and $Q_{\text{cut}} = 20 \text{ GeV}$. The results are shown as the differential jet rates for $1 \rightarrow 0$ (left) and $2 \rightarrow 1$ (right) jets in Figure 7. The solid lines represent the results obtained with the merging algorithm and the dashed lines represent the results obtained without merging algorithms. It is clearly shown that the dependence on the parton shower starting scale is reduced significantly by the merging algorithm.

The inclusive cross section and the ratio of the exclusive cross section to the inclusive cross section are introduced in eqs. (2.25) and (2.26). They are shown in Table 1 with the different scale choices for the parton shower starting scale. Note that the inclusive cross section without merging algorithms is 562 pb. While the inclusive cross section is sensitive to the shower starting scale, the ratios are little affected. This is exactly what we have argued at the end of Section 2.2. The small increase of $\sigma_{\text{exc}}(t\bar{t} + 3)/\sigma_{\text{inc}}$ for the $t_X/4$ may explain the small enhancement at the high k_\perp region in Figure 7.

	$\sigma_{\text{inc}}(t\bar{t})$ (pb)	$\sigma_{\text{exc}}(t\bar{t}+0)/\sigma_{\text{inc}}(t\bar{t})$	$\sigma_{\text{exc}}(t\bar{t}+1)/\sigma_{\text{inc}}(t\bar{t})$	$\sigma_{\text{exc}}(t\bar{t}+2)/\sigma_{\text{inc}}(t\bar{t})$	$\sigma_{\text{exc}}(t\bar{t}+3)/\sigma_{\text{inc}}(t\bar{t})$
s	346	0.28	0.33	0.22	0.17
t_X	422	0.28	0.33	0.22	0.17
$t_X/4$	556	0.27	0.33	0.22	0.18

Table 1: The inclusive cross section and the ratio of the exclusive cross section to the inclusive cross section, with the different choices for the shower starting scale. The inclusive cross section without merging algorithms is 562 pb.

The discussion and the results up to now have assumed that the $t\bar{t}$ is stable. Three jet observables are produced from the event samples including the leptonic decays of the top and antitop quarks and plotted in Figure 8. We use $N = 2$ and $Q_{\text{cut}} = 20$ GeV. The center of mass energy for the pp collisions is 7 TeV. The differential cross section as a function of jet multiplicity is shown in the left panel. The last bin includes the contribution of $N_{\text{jets}} \geq 6$. The gap fraction is defined as

$$f(x) = \frac{N(x)}{N_{\text{total}}}, \quad (2.49)$$

where $N(x)$ is the number of events that give x less than the given value of x . The gap fraction as a function of the p_T of the highest p_T additional jet is shown in the middle panel and that as a function of the scalar sum of the p_T of the additional jets is shown in the right panel. The additional jets are defined as all jets not including the two highest p_T b jets. The green dashed curve shows the result without merging algorithms. The red solid and the blue dotted curves show the results obtained with the merging algorithm. The parton shower starting scale is set to the t_X for the red solid curve and to the $t_X/4$ for the blue dotted curve. Our predictions are compared to the data from the CMS experiment [49]. It is shown that the merging algorithm gives the better description of the data in all the three jet observables. It can also be confirmed that the predictions are stable under the variation of the shower starting scale between the t_X and the $t_X/4$.

Smaller inclusive cross sections obtained with merging algorithms imply less efficient event generation. We therefore choose the scale $t_X/4$ for the parton shower starting scale for generating the event samples to be analyzed in the following sections.

3 Azimuthal angle correlation

In this section, the azimuthal angle difference between the two hardest jets in the top quark pair production process is studied, by using the event samples generated with the merging algorithm described in Section 2. In Section 3.1, the detailed definition of the azimuthal angle difference is specified. In Section 3.2, we discuss the requirements on the merging parameters in order to obtain an accurate prediction of the azimuthal angle difference. A numerical comparison of our merging algorithm with the CKKW-L algorithm is also presented. In Section 3.3, the distribution of the azimuthal angle difference is shown. The result obtained by merging the matrix elements for the $t\bar{t}$ plus up to 2 partons is compared to the one obtained by merging the matrix elements for the $t\bar{t}$ plus up to 3 partons. The roles of the $t\bar{t}+3$ partons matrix elements are studied in detail. It is assumed that the top and antitop quarks are stable in Section 3.3. In Section 3.4, the distribution of the azimuthal angle difference is produced from the event samples including the dilepton decay of the $t\bar{t}$ and the effect of the decay is studied.

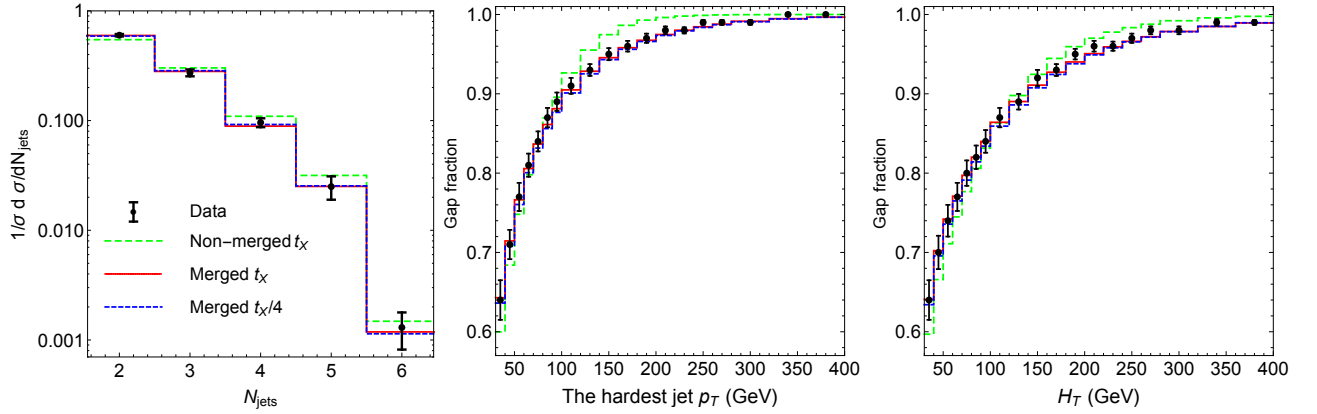


Figure 8: The differential cross section as a function of jet multiplicity (left panel). The gap fraction as a function of the p_T of the highest p_T additional jet (middle panel) and that as a function of the scalar sum of the p_T of the additional jets (right panel). The green dashed curve shows the result without merging algorithms. The red solid and the blue dotted curves show the results obtained with the merging algorithm. The shower starting scale is set to the t_X for the red solid and to the $t_X/4$ for the blue dotted. The merging parameters are set to $N = 2$ and $Q_{\text{cut}} = 20$ GeV. Compared to the data from the CMS experiment [49].

3.1 Definition

An event sample with two or more jets is picked up and the following requirements which are often called vector boson fusion (VBF) cuts are applied to the two hardest jets,

$$y_1 \times y_2 < 0, \quad |y_1 - y_2| > 4. \quad (3.1)$$

The transverse momentum p_T with respect to the beam of an object describes the hardness of the object. Therefore a jet which has the highest p_T is called the hardest jet and another jet which has the second highest p_T is called the second hardest jet, and these jets are assigned to the two hardest jets. One of the two jets which has a positive rapidity is chosen for an azimuthal angle ϕ_1 and the other jet which has a negative rapidity is chosen for an azimuthal angle ϕ_2 . The azimuthal angle difference between the two hardest jets is defined by

$$\Delta\phi = \phi_1 - \phi_2. \quad (3.2)$$

Therefore a jet for the ϕ_1 is not necessarily the hardest jet in our definition. To enhance the correlation in $\Delta\phi$, a cut is applied on the invariant mass of the top quark pair [13],

$$m_{t\bar{t}} < 500 \text{ GeV}. \quad (3.3)$$

No other cuts are applied to the top and anti-top quarks.

All the event samples analyzed in the following sections satisfy the above cuts in eqs. (3.1) and (3.3), even when it is not stated explicitly. This is also the case, when we analyze exclusive parton level event samples in Section 3.3. There, the cuts in eq. (3.1) are applied on partons, not jets.

3.2 Contamination

In this section, we discuss the requirements on the merging parameters in order to obtain an accurate prediction on the azimuthal angle difference. We explore a relation between the merging scale Q_{cut} and a

$p_{T\text{cut}}$ (GeV)	20	25	30
CKKW-L	12.13	7.00	4.50
CKKW-L+	11.65	6.70	4.29

Table 2: The contamination (%) for the two different merging algorithms and for the three different $p_{T\text{cut}}$ values, defined in eq. (3.6). The rapidity cut on jets is $|y| < 4.5$. The merging parameters are set to $N = 3$ and $Q_{\text{cut}} = 20$ GeV.

lower cut on the p_T of jets by which the contamination is negligible. A numerical comparison of our merging algorithm with the CKKW-L algorithm is also presented.

In order for each of the two hardest jets to have the correct azimuthal angle information, each of them must have its origin in a parton generated with matrix elements. If one or both of them originate from partons generated with parton showers, angular correlations between them are not correct. One of the requirements on the merging parameters is therefore

$$N \geq 2. \quad (3.4)$$

Another requirement is that the merging scale Q_{cut} is chosen smaller than the scale of a jet definition. The anti- k_T algorithm [47] includes two parameters, namely the radius parameter R and a lower cutoff on the transverse momentum of jets $p_{T\text{cut}}$. The radius parameter will not be restricted so much, since the two hardest jets which are well separated with each other are of interest to us, see eq. (3.1). We choose $R = 1$ in the merging scale definition in eq. (2.44), while $R = 0.4$ is used in the jet definition. The $p_{T\text{cut}}$ has to satisfy [28]

$$Q_{\text{cut}} \leq p_{T\text{cut}}. \quad (3.5)$$

As we have already mentioned several times, the contamination is the contribution from the $t\bar{t} + 0, 1$ parton matrix elements to the event samples with two or more jets. The contamination can be written by using the inclusive cross section and the exclusive cross sections as

$$\frac{\sigma_{\text{exc}}(t\bar{t} + 0) + \sigma_{\text{exc}}(t\bar{t} + 1)}{\sigma_{\text{inc}}(t\bar{t})}. \quad (3.6)$$

This notation is introduced in eq. (2.25). The $\Delta\phi$ prediction will not be reliable, unless the contamination is small. The event samples are generated with the two different merging algorithms, namely the CKKW-L and the CKKW-L+. Then, the contamination is calculated for the three different $p_{T\text{cut}}$ values, namely 20, 25 and 30 GeV. The rapidity cut on jets is set to $|y| < 4.5$. We use $N = 3$ and $Q_{\text{cut}} = 20$ GeV, thus the above two requirements are satisfied. The result is shown in units of percentage in Table 2. The contamination decreases with a rise in the $p_{T\text{cut}}$, as expected. However, the contamination is not so suppressed in the CKKW-L+ algorithm, compared to the CKKW-L algorithm.

Our introduction of the CKKW-L+ algorithm has been motivated by our numerical finding that the contamination is not negligible in the CKKW-L algorithm when the Q_{cut} is set equal to or slightly smaller than the $p_{T\text{cut}}$. However, it is found that a large suppression of the contamination cannot be achieved in the CKKW-L+ algorithm.

In the following sections, we carry on the further analyses on the event samples generated with the CKKW-L+ algorithm. We choose $Q_{\perp\text{cut}} = 20$ GeV and $p_{T\text{cut}} = 30$ GeV, in order to suppress the contamination reasonably while avoiding the event generation which is too inefficient. Note that about 4% of the correlation in $\Delta\phi$ is already lost due to the contamination in the results presented in the following sections.

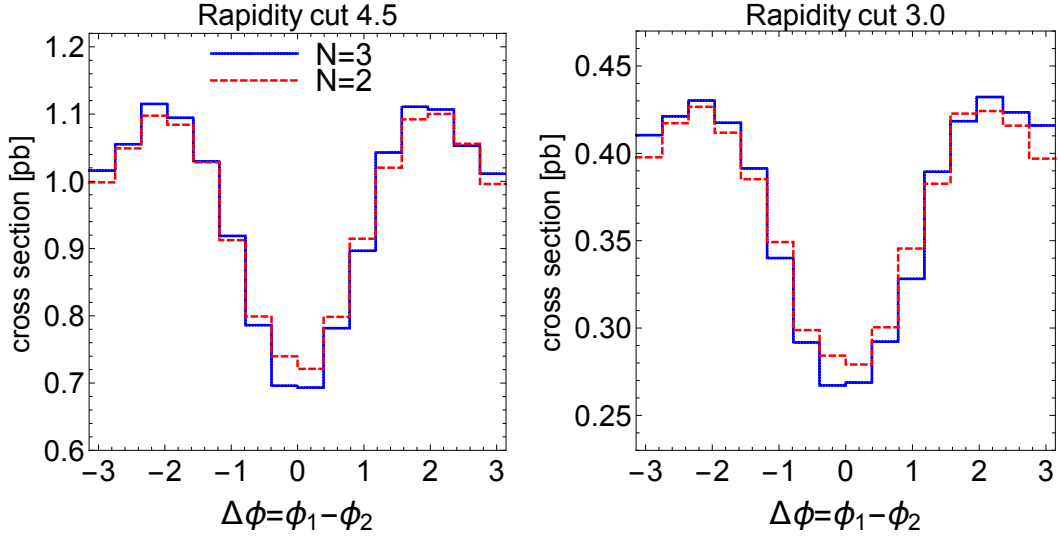


Figure 9: The $\Delta\phi$ distribution produced from the event samples generated with the merging algorithm for $N = 2$ (red dotted curve) and from those for $N = 3$ (blue solid curve). The rapidity cut on jets is $|y| < 4.5$ in the left panel and $|y| < 3.0$ in the right panel.

3.3 The effects of the $t\bar{t}+3$ partons matrix elements

The requirement on the maximal number of partons N provided by the tree level matrix elements (MEs) i.e. $N \geq 2$ is briefly discussed in Section 3.2. In general the more accurate description of multi-jet processes is expected by larger values for N . Thus, the prediction on the $\Delta\phi$ distribution is also expected to be more accurate for $N = 3$ than for $N = 2$. In this section, we produce the $\Delta\phi$ distribution by using the event samples generated with the merging algorithm for $N = 2$ and by using those for $N = 3$. Then, the two results are compared. The effects of the $t\bar{t} + 3$ partons matrix elements on $\Delta\phi$ are studied in detail.

We show the $\Delta\phi$ distribution produced from the event samples generated with the merging algorithm for $N = 2$ (red dotted curve) and from those for $N = 3$ (blue solid curve) in Figure 9. The rapidity cut on jets is $|y| < 4.5$ in the left panel and $|y| < 3.0$ in the right panel. The total inclusive cross section for $m_t = 172.5$ GeV is estimated to be 960pb from ref. [50]. From this value, the cross section at each bin is calculated. The results show strong correlations in $\Delta\phi$, as predicted in the analysis based on the $t\bar{t} + 2$ partons tree level matrix elements [13]. This observation indicates that the correlation found in the $t\bar{t} + 2$ partons tree level matrix elements [13] is still present after including the dominant higher order corrections and thus can be observed in the experiments. We can find a clear difference in the $\Delta\phi$ distribution between the result for $N = 2$ and the one for $N = 3$. The difference looks slightly larger in the right panel, where the rapidity range is more restricted. The origins of the difference are studied in the following.

First of all, let us remind us of one point in the merging algorithm. When the parton shower (PS) generator is executed on a matrix element (ME) event sample for the highest parton multiplicity N i.e. $\{p\}_{t\bar{t}+N}$, the shower is constrained to be softer than the N partons of the $\{p\}_{t\bar{t}+N}$ in terms of the shower evolution variable. More precisely, following a parton shower history of the $\{p\}_{t\bar{t}+N}$ consisting of $\{p\}_{t\bar{t}+(N-1)}, \{p\}_{t\bar{t}+(N-2)}, \dots, \{p\}_{t\bar{t}+i}, \dots, \{p\}_{t\bar{t}+1}, \{p\}_{t\bar{t}}$ with the corresponding scales $t_N < t_{N-1} < \dots < t_{i+1} < \dots < t_2 < t_1$, the evolution scale of the shower is restricted to be below the t_N .

Now let us consider the merging algorithm for $N = 2$ and suppose that the shower evolution of a ME event sample $\{p\}_{t\bar{t}+2}$ is performed and thus an event sample $\{p\}_{t\bar{t}+3}$ is generated. When the shower evolution generates an initial state radiation, one parton will be added in the $\{p\}_{t\bar{t}+2}$ and accordingly the kinematics of the two partons of the $\{p\}_{t\bar{t}+2}$ will be changed. In the left panel of Figure 10, the

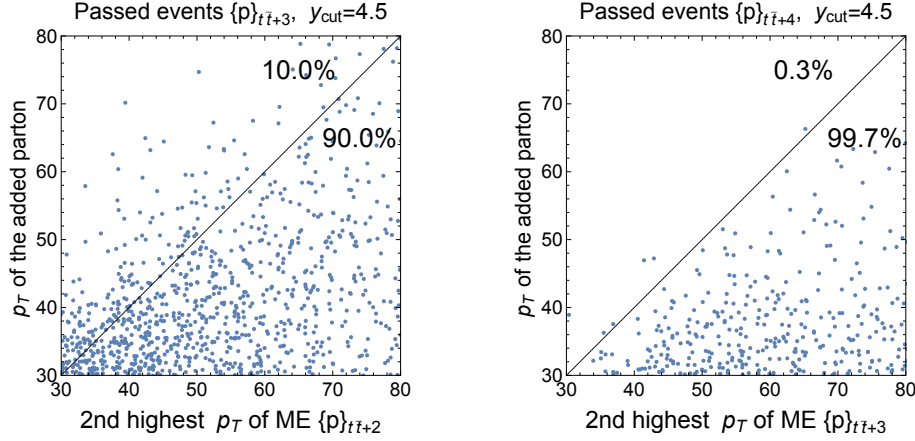


Figure 10: *left*: Event samples $\{p\}_{t\bar{t}+3}$ obtained by one initial state shower evolution of the ME samples $\{p\}_{t\bar{t}+2}$ are plotted. The vertical axis represents p_T of the added parton. The horizontal axis represents the p_T of the second hardest parton of the $\{p\}_{t\bar{t}+2}$ after the evolution. *right*: Event samples $\{p\}_{t\bar{t}+4}$ obtained by one initial state shower evolution of the ME samples $\{p\}_{t\bar{t}+3}$ are plotted. The vertical axis represents p_T of the added parton. The horizontal axis represents the p_T of the second hardest parton of the $\{p\}_{t\bar{t}+3}$ after the evolution.

p_T of the added parton by the initial state radiation is assigned on the vertical axis and the p_T of the second hardest parton of the $\{p\}_{t\bar{t}+2}$ after the initial state radiation is assigned on the horizontal axis. It must be noted that the kinematics of the $\{p\}_{t\bar{t}+2}$ after the initial state radiation is not identical to that of the original ME event sample $\{p\}_{t\bar{t}+2}$ anymore, since it has been changed by the radiation. The rapidity range for the two partons which construct $\Delta\phi$ is $|y| < 4.5$. It is not required, however, that all of the three partons in the $\{p\}_{t\bar{t}+3}$ are within the range $|y| < 4.5$. From the left panel, we can find that the added parton does not have a lower p_T than the two partons of the $\{p\}_{t\bar{t}+2}$ in the considerable fraction (10.0%), which is shown by the dots in the upper left of the panel. This observation implies a non-negligible loss of the correlation between the two hardest jets, because the added parton which has the highest or second highest p_T may give rise to one of the two hardest jets at the end of the shower evolution.

The above problem can be solved by merging the matrix elements for the $t\bar{t}+3$ partons. Let us consider the merging algorithm for $N = 3$ and suppose that the shower evolution of a ME event sample $\{p\}_{t\bar{t}+3}$ is performed and thus an event sample $\{p\}_{t\bar{t}+4}$ is generated. When the shower evolution generates an initial state radiation, one parton will be added in the $\{p\}_{t\bar{t}+3}$ and accordingly the kinematics of the three partons of the original $\{p\}_{t\bar{t}+3}$ will be changed. In the right panel of Figure 10, the p_T of the added parton by the initial state radiation is assigned on the vertical axis and the p_T of the second hardest parton of the $\{p\}_{t\bar{t}+3}$ after the initial state radiation is assigned on the horizontal axis. The rapidity range is the same as above. The panel shows that the probability that the added parton has the highest or second highest p_T is quite suppressed (0.3%).

Our discussion so far has been based on the exclusive parton level event samples, not on the inclusive jet level event samples. However, we believe that our numerical findings reasonably explain the difference observed in Figure 9. In the event samples generated with the merging algorithm for $N = 2$, a non-negligible loss (5 ~ 10%) of the correlation between the two hardest jets can be unavoidable, because a jet originating from hard parton showers can have a higher p_T than one of the two jets originating from the two hard partons of the $t\bar{t}+2$ partons matrix elements. In the event samples generated with the merging algorithm for $N = 3$, the loss of the correlation due to the above reason can be avoided as follows. Jets originating from hard parton showers cannot have higher p_T than two of the three jets originating from the three hard partons of the $t\bar{t}+3$ partons matrix elements.

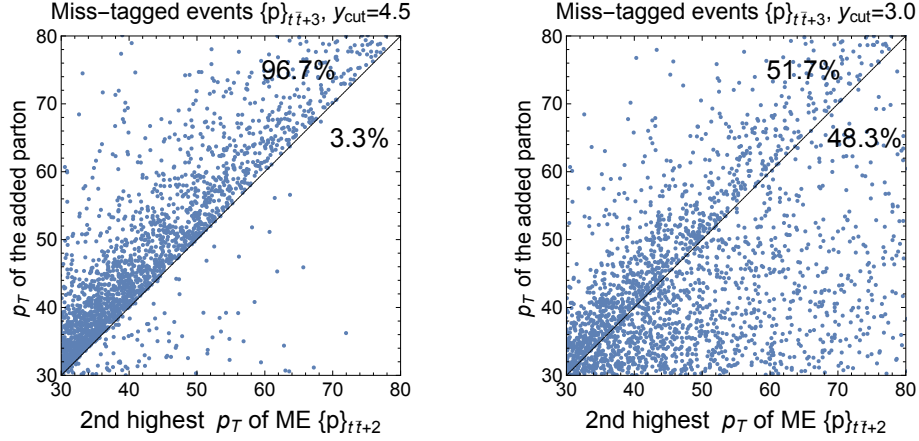


Figure 11: As in the merging algorithm for $N = 2$, event samples $\{p\}_{t\bar{t}+3}$ are generated at first by the initial state shower evolution of the ME event samples $\{p\}_{t\bar{t}+2}$. And then only the miss-tagged event samples are plotted. The axes are identical to those in the left panel of Figure 10. The rapidity range for the two partons which construct $\Delta\phi$ is $|y| < 4.5$ in the left panel and $|y| < 3.0$ in the right panel.

When the rapidity cut on jets is more stringent, a further loss of the correlation may arise, because jets originating from the hard partons of matrix elements may be removed by the rapidity cut. In order to examine the possibility of this, let us define a miss-tagged event sample as an event sample in which a parton added by parton showers or a jet originating from parton showers is picked up for constructing the azimuthal angle difference $\Delta\phi$. In miss-tagged event samples, the $\Delta\phi$ distribution will not be produced correctly, of course.

As in the merging algorithm for $N = 2$, event samples $\{p\}_{t\bar{t}+3}$ are generated by the initial state shower evolution of the ME event samples $\{p\}_{t\bar{t}+2}$. Then, only the miss-tagged event samples are plotted in the left panel of Figure 11. The rapidity range for partons is $|y| < 4.5$. It must be noted that the rapidity range $|y| < 4.5$ is for the two partons which construct $\Delta\phi$ and thus it is not required that all of the three partons in the $\{p\}_{t\bar{t}+3}$ are within the range $|y| < 4.5$. The vertical and horizontal axes are identical to those in the left panel of Figure 10. The miss-tagged event samples in the upper left of the panel correspond to the samples in which the added parton has a higher p_T than one of the two partons of the $\{p\}_{t\bar{t}+2}$ and thus the added parton is picked up for $\Delta\phi$. The miss-tagged event samples in the lower right of the panel correspond to the samples in which one of the two partons of the $\{p\}_{t\bar{t}+2}$ is removed by the rapidity cut and thus the added parton is picked up for $\Delta\phi$, despite that the added parton has a lower p_T . The panel shows that the former possibility is dominant (96.7%). The miss-tagged fraction in the samples $\{p\}_{t\bar{t}+3}$ is 10.3%.

The miss-tagged event samples $\{p\}_{t\bar{t}+3}$ with a more restricted rapidity range $|y| < 3.0$ are plotted in the right panel of Figure 11. The panel shows that one of the two partons of the $\{p\}_{t\bar{t}+2}$ is removed by the rapidity cut in the larger fraction (48.3%) of the samples, as expected. As a result, the miss-tagged fraction in the samples $\{p\}_{t\bar{t}+3}$ is increased to 16.6%.

In event samples $\{p\}_{t\bar{t}+4}$ generated by the initial state shower evolution of the ME event samples $\{p\}_{t\bar{t}+3}$ as in the merging algorithm for $N = 3$, the miss-tagged fraction is 0.4% for a rapidity cut $|y| < 4.5$ and 2.0% for a rapidity cut $|y| < 3.0$. The miss-tagged fraction in the $\{p\}_{t\bar{t}+3}$ is more increased, compared to the $\{p\}_{t\bar{t}+4}$, as the rapidity cut is set more stringent.

Although the above discussion is again based on the exclusive parton level event samples, we believe that our findings reasonably explain the slightly larger difference observed in the right panel of Figure 9. In the event samples generated with the merging algorithm for $N = 2$, a further loss of the correlation between the

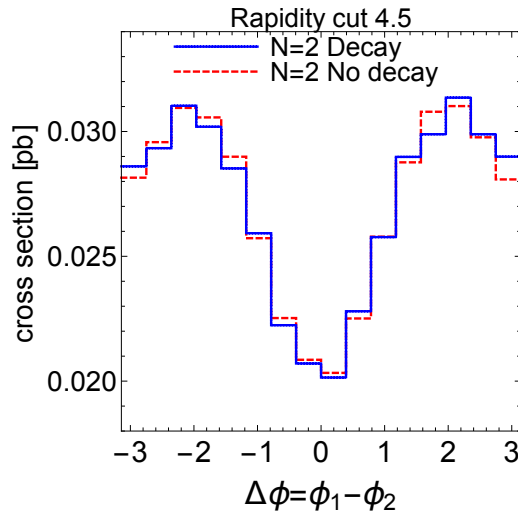


Figure 12: The $\Delta\phi$ distribution produced from the event samples generated by the merging algorithm for $N = 2$, including the dilepton $t\bar{t}$ decay (blue solid) and not including $t\bar{t}$ decays (red dotted).

hardest two jets arises when the rapidity cut on jets is set stringent, because one of the two jets originating from the two hard partons of the $t\bar{t}+2$ partons matrix elements can be removed by the rapidity cut. In the event samples generated with the merging algorithm for $N = 3$, the loss of the correlation due to the above reason can be reduced as follows. Even though one of the three jets originating from the three hard partons of the $t\bar{t}+3$ partons matrix elements is removed by the rapidity cut, there are still two jets which correctly predict the correlation. When the rapidity range for jets is $|y| < 4.5$, the loss of the correlation due to the above reason can be negligible. However, when the rapidity range is more restricted such as $|y| < 3.0$, the loss of the correlation cannot be negligible anymore in the event samples generated with the merging algorithm for $N = 2$, while it can be much reduced in the event samples generated with the merging algorithm for $N = 3$. This observation can explain the larger difference in the right panel of Figure 9.

To summarize, a non-negligible loss of the correlation between the two hardest jets will be unavoidable in the event samples generated with the merging algorithm for $N = 2$. The loss of the correlation can be reduced significantly in the event samples generated with the merging algorithm for $N = 3$. The role of the $t\bar{t}+3$ partons matrix elements can be more important as the rapidity range for jets is set more restricted.

3.4 The effect of a $t\bar{t}$ decay

In the studies of the previous section, it is assumed that the top and antitop quarks are stable. This assumption can be justified, since the purposes of the previous section are to confirm that the correlation in $\Delta\phi$ found in the $t\bar{t}+2$ partons matrix elements [13] is still visible after including the dominant QCD higher order corrections, and to study the effects of the $t\bar{t}+3$ partons matrix elements on $\Delta\phi$. In this section, the effect of a $t\bar{t}$ decay on $\Delta\phi$ is studied.

We produce the $\Delta\phi$ distribution from the event samples including the dilepton $t\bar{t}$ decay generated by the merging algorithm for $N = 2$. Following ref. [49], the following kinematic constraints are imposed. Two oppositely charged leptons (muon or electron) are required to have $p_T > 20$ GeV within the rapidity range $|y| < 2.4$. Jets are rejected if the selected leptons are within a cone of $\Delta R = 0.4$ with respect to the jet. A jet is identified as a b jet if it contains at least one b quark. At least two b jets which fulfill $p_T > 30$ GeV and $|y| < 2.4$ are required. The additional jets are defined as all jets not including the two highest p_T b jets. The two hardest jets for $\Delta\phi$ are picked up from the additional jets. We set the rapidity range for the additional jets as $|y| < 4.5$. In our result, the $t\bar{t}$ invariant mass is calculated from the kinematic information

of the top and antitop quarks obtained just before their decays.

The result is shown by the blue solid curve in Figure 12. The red dotted curve is the result without $t\bar{t}$ decays and is given for comparison. Note that the cross section for the result without $t\bar{t}$ decays is set equal to the one for the result with the $t\bar{t}$ decay. The figure shows that the effect of the $t\bar{t}$ decay on $\Delta\phi$ is small and the correlation is still visible.

4 Conclusion

In this work, the azimuthal angle difference between the two hardest jets (i.e. the two highest p_T jets), $\Delta\phi = \phi_1 - \phi_2$ in the top quark pair production process at the 14 TeV LHC has been studied. The event samples are generated by merging the tree level matrix elements for the $t\bar{t}$ plus up to 2 or 3 partons with the parton shower model in PYTHIA8.

As tree level merging algorithms, we have implemented the CKKW-L algorithm and a new algorithm. Our new algorithm differs from the CKKW-L algorithm in the strategy for phase space separation and it is designed so that the contribution from the $t\bar{t} + 0,1$ parton matrix elements to the event samples with two or more jets, which we call the contamination, can be more suppressed above the merging scale. Although it has been confirmed that the contamination is more suppressed in our algorithm by numerically comparing the two algorithms, the difference is found not drastic. We find that the contamination is not negligible when the merging scale is set equal to or slightly smaller than the scale of the anti- k_T jet definition.

The $\Delta\phi$ distribution is produced from the generated event samples. The distribution shows a strong correlation in $\Delta\phi$, as predicted in the previous analysis [13] based on the $t\bar{t} + 2$ partons tree level matrix elements. This observation confirms that the correlation found in the $t\bar{t} + 2$ partons tree level matrix elements [13] is still visible after including the dominant QCD higher order corrections and thus can be observed in the experiments.

We find a clear difference in the $\Delta\phi$ distribution between the event samples generated by merging the tree level matrix elements for the $t\bar{t}$ plus up to 2 partons and those generated by merging the tree level matrix elements for the $t\bar{t}$ plus up to 3 partons. Furthermore, the difference is found slightly larger, when the rapidity range for jets is more restricted. We have studied the origins of the difference, or in other words the effects of the $t\bar{t} + 3$ partons matrix elements on $\Delta\phi$. When the matrix elements for the $t\bar{t}$ plus up to 2 partons are merged, a non-negligible fraction (5 – 10%) of the correlation between the two hardest jets can be lost because a jet originating from hard parton showers can have a higher p_T than one of the two jets originating from the two hard partons of the $t\bar{t} + 2$ partons matrix elements. When the rapidity range for jets is more restricted, a further loss of the correlation arises because one of the two hard jets can be removed by the rapidity cut. When the matrix elements for the $t\bar{t}$ plus up to 3 partons are merged, the loss of the correlation due to the above two reasons can be avoided as follows. At first, jets originating from hard parton showers cannot have higher p_T than two of the three jets originating from the three hard partons of the $t\bar{t} + 3$ partons matrix elements. Second, even though one of the three hard jets is removed by the rapidity cut, there are still two jets which correctly predict the correlation. Therefore, it can be concluded that, the $t\bar{t} + 3$ partons matrix elements play important roles in predicting $\Delta\phi$ accurately, since they effectively reduce the loss of the correlation. They can be more important as the rapidity cut on jets is set more stringent.

We present a method for merging the matrix element event samples which include a $t\bar{t}$ decay as a part of the hard process with the parton shower. The effect of the $t\bar{t}$ dilepton decay on $\Delta\phi$ is studied. We have shown that the effect is small when the two hardest jets for $\Delta\phi$ are picked up from all jets not including the two hardest b jets.

We note that our findings should be applicable equally to other heavy particle production processes by gluon fusion. We hope that our findings help experimentalists perform the proposal of ref. [13] and achieve

precise measurement of the CP property of the Higgs boson by using the azimuthal angle correlation between the two hardest jets at the LHC.

Acknowledgments

JN is grateful to Rikkert Frederix for many helpful discussions and his encouragement, to Junichi Kanzaki for giving valuable comments on the plots and to Yasuhito Sakaki for fruitful discussions on simulating decays of a top quark. JN would also like to acknowledge the warm hospitality of the Theory Division at CERN, where the first step of the present work was carried out. The authors are grateful to KEK Computing Research Center for the management of KEKCC on which most of the simulation in this study was executed. This work is supported in part by JSPS KAKENHI Grant No. 25.4461 of JN.

References

- [1] G. Aad *et al.* [ATLAS Collaboration], “Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC,” *Phys. Lett. B* **726** (2013) 88 [Erratum-ibid. *B* **734** (2014) 406] [arXiv:1307.1427 [hep-ex]].
- [2] G. Aad *et al.* [ATLAS Collaboration], “Measurements of fiducial and differential cross sections for Higgs boson production in the diphoton decay channel at $\sqrt{s} = 8$ TeV with ATLAS,” *JHEP* **1409** (2014) 112 [arXiv:1407.4222 [hep-ex]].
- [3] S. Chatrchyan *et al.* [CMS Collaboration], “Measurement of the properties of a Higgs boson in the four-lepton final state,” *Phys. Rev. D* **89** (2014) 092007 [arXiv:1312.5353 [hep-ex]].
- [4] V. Khachatryan *et al.* [CMS Collaboration], “Observation of the diphoton decay of the Higgs boson and measurement of its properties,” *Eur. Phys. J. C* **74** (2014) 10, 3076 [arXiv:1407.0558 [hep-ex]].
- [5] T. Plehn, D. L. Rainwater and D. Zeppenfeld, “Determining the structure of Higgs couplings at the LHC,” *Phys. Rev. Lett.* **88** (2002) 051801 [hep-ph/0105325].
- [6] V. Del Duca, W. Kilgore, C. Oleari, C. Schmidt and D. Zeppenfeld, “Gluon fusion contributions to $H + 2$ jet production,” *Nucl. Phys. B* **616** (2001) 367 [hep-ph/0108030].
- [7] G. Klamke and D. Zeppenfeld, “Higgs plus two jet production via gluon fusion as a signal at the CERN LHC,” *JHEP* **0704** (2007) 052 [hep-ph/0703202 [HEP-PH]].
- [8] K. Hagiwara, Q. Li and K. Mawatari, “Jet angular correlation in vector-boson fusion processes at hadron colliders,” *JHEP* **0907** (2009) 101 [arXiv:0905.4314 [hep-ph]].
- [9] V. Del Duca, G. Klamke, D. Zeppenfeld, M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau and A. D. Polosa, “Monte Carlo studies of the jet activity in Higgs + 2 jet events,” *JHEP* **0610** (2006) 016 [hep-ph/0608158].
- [10] C. Ruwiedel, N. Wermes and M. Schumacher, “Prospects for the measurement of the structure of the coupling of a Higgs boson to weak gauge bosons in weak boson fusion with the ATLAS detector,” *Eur. Phys. J. C* **51** (2007) 385.
- [11] J. R. Andersen, V. Del Duca and C. D. White, “Higgs Boson Production in Association with Multiple Hard Jets,” *JHEP* **0902** (2009) 015 [arXiv:0808.3696 [hep-ph]].
- [12] J. R. Andersen, K. Arnold and D. Zeppenfeld, “Azimuthal Angle Correlations for Higgs Boson plus Multi-Jet Events,” *JHEP* **1006** (2010) 091 [arXiv:1001.3822 [hep-ph]].
- [13] K. Hagiwara and S. Mukhopadhyay, “Azimuthal correlation among jets produced in association with a bottom or top quark pair at the LHC,” *JHEP* **1305**, 019 (2013) [arXiv:1302.0960 [hep-ph]].

- [14] S. Catani, F. Krauss, R. Kuhn and B. R. Webber, “QCD matrix elements + parton showers,” JHEP **0111**, 063 (2001) [hep-ph/0109231].
- [15] F. Krauss, “Matrix elements and parton showers in hadronic interactions,” JHEP **0208**, 015 (2002) [hep-ph/0205283].
- [16] S. Hoeche, F. Krauss, S. Schumann and F. Siegert, “QCD matrix elements and truncated showers,” JHEP **0905** (2009) 053 [arXiv:0903.1219 [hep-ph]].
- [17] K. Hamilton, P. Richardson and J. Tully, “A Modified CKKW matrix element merging approach to angular-ordered parton showers,” JHEP **0911** (2009) 038 [arXiv:0905.3072 [hep-ph]].
- [18] L. Lonnblad, “Correcting the color dipole cascade model with fixed order matrix elements,” JHEP **0205** (2002) 046 [hep-ph/0112284].
- [19] N. Lavesson and L. Lonnblad, “W+jets matrix elements and the dipole cascade,” JHEP **0507** (2005) 054 [hep-ph/0503293].
- [20] L. Lonnblad and S. Prestel, “Matching Tree-Level Matrix Elements with Interleaved Showers,” JHEP **1203** (2012) 019 [arXiv:1109.4829 [hep-ph]].
- [21] W. T. Giele, D. A. Kosower and P. Z. Skands, “Higher-Order Corrections to Timelike Jets,” Phys. Rev. D **84**, 054003 (2011) [arXiv:1102.2126 [hep-ph]].
- [22] M. L. Mangano, M. Moretti and R. Pittau, “Multijet matrix elements and shower evolution in hadronic collisions: $Wb\bar{b} + n$ jets as a case study,” Nucl. Phys. B **632** (2002) 343 [hep-ph/0108069].
- [23] M. L. Mangano, M. Moretti, F. Piccinini and M. Treccani, “Matching matrix elements and shower evolution for top-quark production in hadronic collisions,” JHEP **0701** (2007) 013 [hep-ph/0611129].
- [24] S. Mrenna and P. Richardson, “Matching matrix elements and parton showers with HERWIG and PYTHIA,” JHEP **0405** (2004) 040 [hep-ph/0312274].
- [25] J. Alwall, S. de Visscher and F. Maltoni, “QCD radiation in the production of heavy colored particles at the LHC,” JHEP **0902** (2009) 017 [arXiv:0810.5350 [hep-ph]].
- [26] J. Alwall, S. Hoeche, F. Krauss, N. Lavesson, L. Lonnblad, F. Maltoni, M. L. Mangano and M. Moretti *et al.*, “Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions,” Eur. Phys. J. C **53** (2008) 473 [arXiv:0706.2569 [hep-ph]].
- [27] N. Lavesson and L. Lonnblad, “Merging parton showers and matrix elements: Back to basics,” JHEP **0804** (2008) 085 [arXiv:0712.2966 [hep-ph]].
- [28] M. R. Buckley, T. Plehn and M. J. Ramsey-Musolf, “Top squark with mass close to the top quark,” Phys. Rev. D **90** (2014) 1, 014046 [arXiv:1403.2726 [hep-ph]].
- [29] S. Mukhopadhyay, M. M. Nojiri and T. T. Yanagida, “Compressed SUSY search at the 13 TeV LHC using kinematic correlations and structure of ISR jets,” JHEP **1410** (2014) 12 [arXiv:1403.6028 [hep-ph]].
- [30] M. Buschmann, C. Englert, D. Goncalves, T. Plehn and M. Spannowsky, “Resolving the Higgs-Gluon Coupling with Jets,” Phys. Rev. D **90** (2014) 013010 [arXiv:1405.7651 [hep-ph]].
- [31] M. Buschmann, D. Goncalves, S. Kuttimalai, M. Schonherr, F. Krauss and T. Plehn, “Mass Effects in the Higgs-Gluon Coupling: Boosted vs Off-Shell Production,” arXiv:1410.5806 [hep-ph].
- [32] V. N. Gribov and L. N. Lipatov, “Deep inelastic e-p scattering in perturbation theory, Sov. J. Nucl. Phys. **15** (1972) 438.
- [33] G. Altarelli and G. Parisi, “Asymptotic freedom in parton language, Nucl. Phys. B **126** (1977) 298318.

- [34] Y. L. Dokshitzer, “Calculation of the structure functions for deep inelastic scattering and e^+e^- annihilation by perturbation theory in quantum chromodynamics, Sov. Phys. JETP **46** (1977) 641653.
- [35] G. Marchesini and B. R. Webber, “Monte Carlo Simulation of General Hard Processes with Coherent QCD Radiation,” Nucl. Phys. B **310**, 461 (1988).
- [36] T. Sjostrand, “A Model for Initial State Parton Showers,” Phys. Lett. B **157** (1985) 321.
- [37] T. D. Gottschalk, “Backwards Evolved Initial State Parton Showers,” Nucl. Phys. B **277**, 700 (1986).
- [38] W. T. Giele, D. A. Kosower and P. Z. Skands, “A simple shower and matching algorithm,” Phys. Rev. D **78**, 014026 (2008) [arXiv:0707.3652 [hep-ph]].
- [39] S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, “Longitudinally invariant K_t clustering algorithms for hadron hadron collisions,” Nucl. Phys. B **406** (1993) 187.
- [40] T. Sjostrand and P. Z. Skands, “Transverse-momentum-ordered showers and interleaved multiple interactions,” Eur. Phys. J. C **39** (2005) 129 [hep-ph/0408302].
- [41] R. Corke and T. Sjostrand, “Interleaved Parton Showers and Tuning Prospects,” JHEP **1103** (2011) 032 [arXiv:1011.1759 [hep-ph]].
- [42] E. Norrbin and T. Sjostrand, “QCD radiation off heavy particles,” Nucl. Phys. B **603** (2001) 297 [hep-ph/0010012].
- [43] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao and T. Stelzer *et al.*, “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations,” JHEP **1407**, 079 (2014) [arXiv:1405.0301 [hep-ph]].
- [44] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. M. Nadolsky and W. K. Tung, “New generation of parton distributions with uncertainties from global QCD analysis,” JHEP **0207** (2002) 012 [hep-ph/0201195].
- [45] T. Sjostrand, S. Mrenna and P. Z. Skands, “A Brief Introduction to PYTHIA 8.1,” Comput. Phys. Commun. **178** (2008) 852 [arXiv:0710.3820 [hep-ph]].
- [46] T. Sjostrand, S. Mrenna and P. Z. Skands, “PYTHIA 6.4 Physics and Manual,” JHEP **0605** (2006) 026 [hep-ph/0603175].
- [47] M. Cacciari, G. P. Salam and G. Soyez, “The Anti- $k(t)$ jet clustering algorithm,” JHEP **0804** (2008) 063 [arXiv:0802.1189 [hep-ph]].
- [48] M. Cacciari, G. P. Salam and G. Soyez, “FastJet User Manual,” Eur. Phys. J. C **72** (2012) 1896 [arXiv:1111.6097 [hep-ph]].
- [49] S. Chatrchyan *et al.* [CMS Collaboration], “Measurement of jet multiplicity distributions in $t\bar{t}$ production in pp collisions at $\sqrt{s} = 7$ TeV,” Eur. Phys. J. C **74**, 3014 (2015) [Eur. Phys. J. C **75**, no. 5, 216 (2015)] [arXiv:1404.3171 [hep-ex]].
- [50] M. Czakon, P. Fiedler and A. Mitov, “Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\frac{4}{s})$,” Phys. Rev. Lett. **110**, 252004 (2013) [arXiv:1303.6254 [hep-ph]].